

Bigger uncertainties and the Big Bang

L Tenorio[†], P B Stark[‡] and C H Lineweaver[§]

[†] Mathematical and Computer Science Department, Colorado School of Mines, Golden, CO 80401, USA

[‡] Statistics Department, University of California, Berkeley, CA 94720, USA

[§] Observatoire Astronomique de Strasbourg, 67000 Strasbourg, France

Received 16 September 1998

Abstract. We use Bayesian hierarchical models and recent results from the theory of minimax confidence interval estimation to study the effect of prior information in a cosmological inverse problem. We consider the effect of prior information on uncertainty estimates of a linear functional Lx of an infinite-dimensional model x , given noisy observations $y = Kx + x$. The model represents the cosmic microwave background (CMB), which is the radiation left over from the Big Bang. The linear functional is related to the important cosmological question of whether the CMB temperature varies with direction in the sky; such variation is required by cosmological theories to account for the observed large-scale heterogeneity of matter and energy in the Universe. Evidence of this heterogeneity is the non-zero quadrupole term in the CMB detected by the COBE satellite in 1992. Estimation of the quadrupole is an interesting ill-posed problem that requires more information than cosmologists expected. Previously published quadrupole estimates relied on constraints such as artificially truncating the spherical harmonic expansion of the CMB fluctuations, or on modelling the effect of unestimated high-frequency terms, without accounting for model uncertainty. If these implicit constraints were relaxed, the uncertainty would be several to dozens of times larger than reported in the astrophysical literature. We study the dependence of quadrupole estimates to a series of increasingly stringent constraints. We show that no useful estimates can be obtained from COBE data without assuming a particular class of prior cosmological models. Even restricting the spectrum to lie in a two-parameter family of models commonly used in cosmology does not suffice without positing a prior probability distribution on those two parameters.

1. Introduction

Cosmology is a scientific attempt to understand how the Universe came to be and how it will evolve. Over the past century progress has been made towards answering these questions and has resulted in the well known Big Bang model describing the evolution of the Universe from a primeval explosion. The bath of microwave photons coming from this hot early epoch in the evolution of the Universe is the cosmic microwave background (CMB) radiation. The CMB is made of the oldest photons we can observe, and contains information about the Universe from times much before the birth of galaxies and quasars. The CMB is thus a unique tool for probing the early Universe.

Cosmological models require initial small-density inhomogeneities to explain observed large-scale structures such as galaxies, clusters of galaxies, giant voids and superclusters. These small inhomogeneities were expected to leave their mark as small fluctuations, usually referred to as anisotropies, in the CMB. For 28 years, cosmologists measured the CMB in search of such anisotropies. In 1989 NASA launched its first cosmology satellite, the Cosmic Background Explorer (COBE), carrying the Differential Microwave Radiometer (DMR) instrument on

board to search for CMB anisotropies. In the spring of 1992 the COBE DMR team announced the discovery of anisotropies in the CMB (Smoot *et al* 1992), starting a new era in cosmology (see Silk (1997) for a gentle introduction to cosmology).

The CMB is a function of direction in the sky. To study CMB heterogeneity we look at its spatial frequency components with respect to an orthonormal basis of spherical harmonics on the sphere. In a given direction \hat{r} , the CMB temperature $T(\hat{r})$ is written as

$$T(\hat{r}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{r}) \quad (1)$$

where $Y_{\ell m}$ is the real m th-spherical harmonic of degree ℓ . The first few terms in (1) can be written as: $a_{00} + \langle \mathbf{D}, \hat{r} \rangle + \langle \mathbf{C}\hat{r}, \hat{r} \rangle$, where the vector \mathbf{D} is determined by the *dipole* coefficients $\{a_{1m}\}$, and the symmetric matrix \mathbf{C} by the *quadrupole* coefficients $\{a_{2m}\}$, $-2 \leq m \leq 2$. In the following we shall frequently use the quadrupole coefficients Q_1, \dots, Q_5 ; these are normalized versions of the $\{a_{2,m}\}$ traditionally used in the astrophysical literature (see Kogut *et al* 1996). The quadrupole *principal axes* are the normalized eigenvectors of \mathbf{C} . Since the DMR instrument measured spatial temperature differences in the CMB, DMR data are insensitive to the *monopole* a_{00} . In addition, the dipole is dominated by Earth's motion relative to the CMB, so the quadrupole coefficients are the lowest frequency terms of cosmological significance for DMR data. The quadrupole observed by COBE is the largest detected structure in the Universe and it is evidence of the anisotropic nature of the CMB. (Note, Banday *et al* (1997), Bennett *et al* (1996), Kogut *et al* (1996) and Smoot *et al* (1992) will be abbreviated as Ba97, Be96, K96 and S92 hereafter.)

The data we use are the DMR four-year sky maps. These maps are temperatures in 6144 directions in the sky (pixels) obtained by fitting the time-ordered data of differential measurements corrected for various systematic effects. There are DMR maps for two independent channels measuring the CMB in each of three microwave frequencies: 31 GHz, 53 GHz and 90 GHz. To increase the signal-to-noise ratio we use channel averages of each of the less noisy maps (53 GHz and 90 GHz).

Quadrupole coefficients were originally estimated by least-squares (LS) fitting to sky maps with a 'Galactic cut', i.e. omitting pixels near the Galactic plane to reduce foreground Galactic contamination (S92). More recently, K96 improve quadrupole estimates by including information from power-law models. By a *power-law* model we shall mean a particular parametric class of homogeneous random fields on the sphere used to model the CMB (section 4). These models are characterized by two *spectral* parameters denoted by Q and η . Cosmologists are mainly interested in Q , η and the power spectrum of the model; they believe that no useful science can be derived from the CMB quadrupole in our sky. We consider the quadrupole for the following reasons: (1) assuming a power-law model, the expected value of the quadrupole power $Q_{\text{rms}}^2 (\equiv \sum_{m=-2}^2 a_{2m}^2 / 4\pi)$ is Q^2 . If an estimate of Q_{rms}^2 turns out to be 'very different' from the accepted value of Q^2 , then either there is something wrong with the cosmological model assumptions or we live in a very unusual Universe; (2) the questions we raise, and the methods we use, are also applicable to uncertainty estimates of higher-order harmonic coefficients; (3) a lot of importance was given to the discovery of the quadrupole as evidence of the anisotropic nature of the CMB; and (4) cosmological theories may change to accommodate new observations, the CMB quadrupole in the sky will not.

Standard results from linear inverse theory imply that quadrupole uncertainty estimates from maps with gaps are arbitrarily large unless *some* constraint is imposed on the CMB fluctuations (Stark 1993, Bunn *et al* 1994). K96 use estimates of the spectral parameters to model the expected high ($\ell \geq 3$) spatial frequency contributions; they also use Galactic templates (models) in addition to Galactic cuts to reduce Galactic contamination. Our main

goal is to study the sensitivity of quadrupole estimates to some common implicit constraints, and to assess the prior information required to obtain ‘small’ quadrupole uncertainties. By small we shall mean uncertainties in Q_i less than $10 \mu\text{K}$; this is the order of magnitude of published quadrupole uncertainties.

In section 2 we explore the uncertainty in truncated LS quadrupole estimates. This corresponds to the original estimates in S92, which fitted to an expansion truncated at $\ell_{\text{max}} = 2$. We show that LS uncertainties are extremely sensitive to the order at which the expansion is truncated, and that LS introduces a bias in the direction of the quadrupole axes. By modelling the $\ell \geq 3$ structure, we find uncertainties that are up to three times larger than the LS uncertainties that include only noise variance. This variability, caused by the intrinsic variability of the model and by model uncertainty, is not reflected in previously published estimates.

Having established that more prior information is required to control quadrupole uncertainties, we investigate how strong a constraint we need. In section 3 we use some results on minimax estimation of linear functionals to examine how well the DMR data together with a mild constraint on the CMB allow the control of quadrupole uncertainties without imposing a cosmological model. We find lower bounds for the lengths of quadrupole confidence intervals and find that at least 30 times as many observations per pixel would be required for the quadrupole uncertainties to be as small as those published. We conclude that the usefulness of this root-mean-square (RMS) constraint is limited given the noise level of the DMR maps. We thus go one step further: section 4 incorporates power-law models to constrain the shape of the CMB spectrum in conjunction with the RMS constraint. The model specifies a prior distribution for $\{a_{\ell m}\}$; estimating the quadrupole can be viewed as a typical Bayesian estimation problem except that Q and η are unknown. We use two different methods to control Q, η : an RMS constraint, which turns out to be too weak, and a joint prior distribution. Only in the latter case do we reduce the uncertainties to the desired level.

Section 5 summarizes our conclusions; the appendices contain mathematical and computational details. The monopole and dipole have not been subtracted from the DMR maps we use. Quadrupole components are in Galactic coordinates and have been corrected for the kinematic quadrupole induced by our motion with respect to the CMB (Lineweaver *et al* 1996). We use only Galactic cuts to avoid additional systematic errors that modelling Galactic emissions may introduce; incorporating Galactic models would further increase the uncertainties.

2. Truncated least-squares estimates

The observed temperatures $d = (d_i)$ in the DMR maps are modelled as

$$d_i = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} G_{\ell} a_{\ell m} Y_{\ell m}(\hat{r}_i) + \epsilon_i \quad i = 1, \dots, 6144 \quad (2)$$

where the $a_{\ell m}$ are the coefficients of the true CMB fluctuations. Since DMR data are not sensitive to the monopole, $a_{0,0}$ is just an offset introduced in the map making procedure. To a good approximation the measurement errors ϵ_i are independent, zero-mean Gaussian variables with known variances σ_i^2 (Lineweaver *et al* 1994). The $\{G_{\ell}\}$ are the known Legendre coefficients of the instrument’s beam (Wright *et al* 1994). Because of Galactic foreground contamination, pixels in a range of Galactic latitudes $|b| \leq b_c$ near the Galactic equator $b = 0$ are excluded; typically, b_c is 20° or 20°e . The latter is the extended cut, which discards pixels within 20° of the Galactic plane and pixels in the Orion and Ophiuchus constellations (see Be96). Owing to the Galactic cut, the data alone do not constrain the values of the spherical

harmonic coefficients $\{a_{\ell m}\}$ (Stark 1993, Bunn *et al* 1994); we need to control the component of each $Y_{\ell m}$ to which the data are insensitive.

One standard constraint, implicit in truncated LS fits, is that $\{a_{\ell m}\}$ all vanish beyond some low finite degree $\ell_{\max} \geq 2$. A side effect of this assumption is that the uncertainties in the fitted coefficients strongly depend on ℓ_{\max} . Because standard cosmological models predict that the $\{a_{\ell m}\}$ differ substantially from zero over a larger range of values of ℓ , there is no physical justification for truncating at low degrees.

A second shortcoming, mathematically related to a problem in geophysics (Stark and Hengartner 1993, Pulliam and Stark 1993, Stark 1995), is that fitting a truncated spherical harmonic expansion to noisy data with gaps tends to produce spurious large structure in the gaps. The explanation is simple: the noise and unmodelled part of the real CMB signal have high spatial frequencies, while the model is a linear combination of functions with relatively small spatial derivatives. The LS estimate matches the large derivatives in the data as well as possible. The only way to get large spatial derivatives in a linear combination of smooth functions is by using large coefficients. By Parseval's theorem, if the harmonic coefficients are large, the model itself must be large somewhere. The model can afford to be large, without adversely affecting the fit to the data, in the spatial gaps where there are no data to constrain it.

Figures 1 and 2 illustrate how the principal axes of LS quadrupole estimates tend to align with the Galactic cut. The figures show histograms on the sphere of estimated directions of principal quadrupole axes in the following simulation. Generate a sky map according to a power-law model truncated at $\ell = 25$ (see section 4), with Q and η picked at random according to normal distributions with means and standard deviations based on the four-year data (Be96); namely $Q = 15.3 \pm 3.8$ and $\eta = 1.2 \pm 0.3$. Add independent, DMR-like Gaussian noise to each pixel. Introduce a 20° Galactic cut, and estimate the quadrupole components using LS truncated at $\ell_{\max} = 2$. Repeat 7000 times. Figure 1 shows that when no Galactic cut is used the distribution of the directions is uniform; this is consistent with the cosmological principle of no preferred direction. Figure 2 shows the results when a 20° Galactic cut is used. The quadrupole axes cluster near the Galactic equator, where there are no data, and perpendicular to it (to preserve orthogonality), thus biasing the direction of the quadrupole axes.

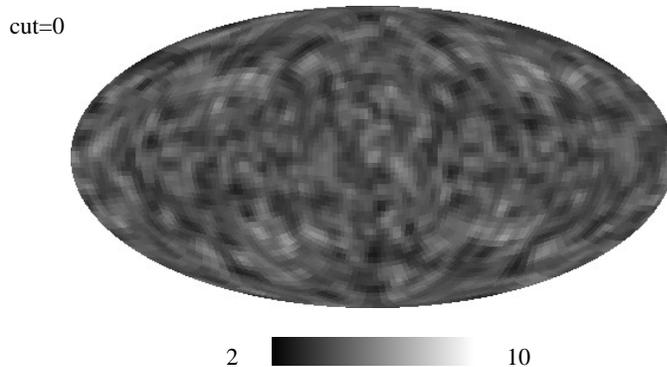


Figure 1. Histogram on the sphere of the directions of quadrupole axes from LS fits with 0° Galactic cut. The map is a planar projection of the sphere, the Galactic plane is along the equator of the map. Each realization adds a count to the three pixels where the axes point.

The quadrupole power is also affected by the Galactic cut. Table 1 presents the RMS departure of each quadrupole estimate from the true value in each simulation, the formal 1σ $\ell_{\max} = 2$ LS uncertainty, and the ratio of the two. Q_{RMS}^2 is estimated using the estimated

cut=20

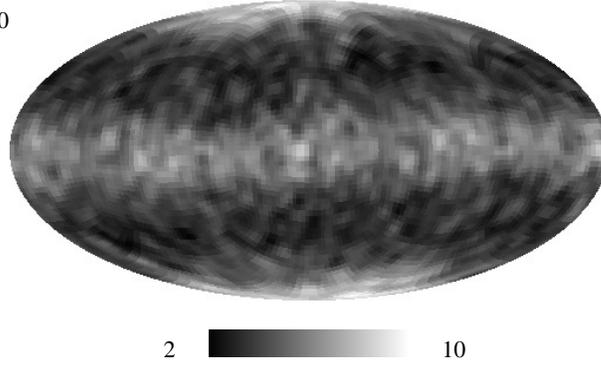


Figure 2. Pulling of quadrupole axes by least-squares. Histogram on the sphere of the directions of quadrupole axes from LS fits with 20° Galactic cut. The distribution is no longer uniform and tends to concentrate along the Galactic plane.

Table 1. Standard errors of LS estimates of quadrupole components (μK), truncated at $\ell_{\text{max}} = 2$, compared with the RMS errors of the truncated LS estimates for random CMB generated from power-law models. The RMS uncertainty from simulations is compared to the formal standard error of the estimator.

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_{RMS}^2
RMS error, no cut	3.0	9.6	10.1	2.5	2.5	44.7
LS σ , no cut	2.0	7.0	6.6	1.7	1.6	2.4
Ratio, no cut	1.5	1.4	1.5	1.5	1.6	18.6
RMS error, 20° cut	8.6	11.4	10.9	8.7	8.6	125.8
LS σ , 20° cut	2.8	7.4	7.0	2.8	2.6	4.8
Ratio, 20° cut	3.1	1.5	1.6	3.1	3.3	26.2

quadrupole coefficients and correcting for the noise bias (K96, Gould 1993). Even without a Galactic cut, the LS uncertainties can be 50% smaller than the RMS error. For the 20° Galactic cut, the uncertainty of Q_{RMS}^2 , measured by the RMS departures in the simulations, is 26 times larger than the formal LS uncertainty. Table 1 shows that LS quadrupole uncertainties are too optimistic: adding information from power-law models and including the modelling uncertainties yields much larger uncertainties in the LS estimates. But what is the best we can do without assuming a cosmological model? We answer this next; we will see that the RMS errors in table 1 are also optimistic.

3. Minimax confidence intervals with RMS constraints

To reduce the aliasing of higher-order multipoles in quadrupole estimates, K96 assume a power-law model for the $\ell \geq 3$ harmonic coefficients and include their covariance as part of the stochastic noise in LS quadrupole fits. This aliasing adjustment is subject to sources of uncertainty not reflected in previously published estimates: (1) possible errors in representing the CMB using any power-law model; (2) errors in the spectral parameters used to estimate the expected aliasing, assuming that the CMB is correctly described by some power-law model; and (3) random deviations of the realized aliasing from the aliasing expected under the model, assuming that the model is correct. We now address (1) by finding lower bounds for the lengths of quadrupole confidence intervals without assuming a cosmological model. The following

sections will address (2) and (3).

The energy in the CMB is finite, we would not be here otherwise, so there is a bound M on the average of the squared, beam-convolved CMB temperature, i.e.

$$\int T^2(\hat{r}) d\Omega = \sum_{\ell,m} G_\ell^2 a_{\ell m}^2 \leq M^2. \quad (3)$$

This mild constraint limits what the CMB can do in the Galactic cut, and allows us to find lower bounds for the lengths of 68% confidence intervals that could be attained for the *best possible estimators* of the quadrupole coefficients, given the data sampling, Galactic cut and noise level. This is a minimax statistical estimation problem; a general theory of finding such lower bounds for linear inverse problems with Gaussian noise was derived in Donoho (1994). Donoho's theory is applied to geophysical problems mathematically related to the present formulation in Stark (1992), Stark and Hengartner (1993) and Pulliam and Stark (1994). Appendix A contains mathematical and computational details relevant to the CMB problem.

We take $M = \sqrt{4\pi}\text{RMS}$ with $\text{RMS} = 35 \mu\text{K}$ (Ba97). This bound was estimated from the four-year data with an error $2 \mu\text{K}$ but we also determine the value of M that yields a lower bound comparable to the length of confidence intervals reported in S92. Table 2 compares the values of the uncertainty bounds, computed for the noise and observation pattern of the 53 GHz sky map, with the LS uncertainty, truncated at $\ell_{\text{max}} = 2$ and 16, for different Galactic cuts. The lower bounds are based on the best possible estimator, so proper confidence intervals around the truncated LS estimate must be at least as long as these bounds. The table shows that when a Galactic cut is used, truncating the spherical harmonic expansion at $\ell_{\text{max}} = 2$ injects much stronger information than the RMS bound, resulting in artificially small uncertainties in LS estimates. The similarities of the bounds for $a_{\ell m}$ and $a_{\ell, -m}$ result from the spatial symmetry of the spherical harmonics. The difference in magnitudes reflects the geometry of the harmonics with respect to the cut, as well as the uneven data sampling. As expected, some of the symmetry is broken by the asymmetric extended cut.

Table 2. Lower bounds on the half-lengths of 68% confidence intervals for individual quadrupole coefficients for the estimator that would admit the shortest possible confidence intervals, incorporating an RMS constraint on the fluctuations (top) and standard errors of LS estimates (bottom). The lower bounds are typically much larger than the formal LS uncertainty for $\ell_{\text{max}} = 2$; truncating LS estimates at degree 2 injects far more information than the RMS bound.

Cut	$a_{2,-2}$	$a_{2,-1}$	$a_{2,0}$	$a_{2,1}$	$a_{2,2}$
Minimax lower bound (μK)					
$\pm 0^\circ$	2	2	2	2	2
$\pm 20^\circ$	93	32	7	31	91
$\pm 20^\circ\text{e}$	97	38	12	31	92
$\sigma_{a_{\ell m}}$ from LS (μK)					
$\pm 20^\circ, \ell_{\text{max}} = 2$	8	5	7	5	7
$\ell_{\text{max}} = 16$	550	27	389	24	384

Transforming to the Q_i normalization, the minimax lower bounds of table 2 set the following lower limits on '1 σ detections' given the noise, observation pattern, 20° Galactic cut and RMS bound, for the best possible estimator (μK):

$$|Q_1| \geq 5 \quad |Q_2| \geq 71 \quad |Q_3| \geq 69 \quad |Q_4| \geq 52 \quad |Q_5| \geq 50.$$

Some of these lower bounds are 10 times larger than the reported detections. For the lower bound on the minimax uncertainty to be as small as the uncertainty in the $\ell_{\text{max}} = 2$ LS estimate, we would need an upper bound M on the RMS of the order of $3 \mu\text{K}$, less than one tenth of the

value estimated by Ba97. Simulations also give the signal-to-noise ratio required to attain the published level of uncertainties: the noise per pixel has to be at least five times smaller than that of DMR data.

4. Power-law models as constraints

We now include information from cosmological models. In large-scale structure theories, the CMB is a realization of a homogeneous random field. We use random fields described by *power-law* models. Under these models the $\{a_{\ell m}\}$, $\ell \geq 2$, are independent $N(0, \sigma_\ell^2)$ variables. The common variance of the $2\ell + 1$ coefficients of degree ℓ is (Bond and Efstathiou 1987)

$$\sigma_\ell^2(Q) = \frac{4\pi Q^2}{5} \frac{\Gamma(\ell + (\eta - 1)/2)\Gamma((9 - \eta)/2)}{\Gamma(\ell + (5 - \eta)/2)\Gamma((3 + \eta)/2)} \quad (4)$$

where $Q \equiv (Q, \eta)$; $Q > 0$, $-1 \leq \eta \leq 2$. The spectrum of the field is the sequence $\{\sigma_\ell^2\}$; a homogeneous Gaussian random field is characterized by its spectrum.

4.1. Prior distributions for the $\ell \leq 2$ terms

For known Q , equation (4) defines a prior joint Gaussian probability distribution for $a_{\ell m}$, $\ell \geq 2$. In this section we compute posterior conditional estimates of the quadrupole coefficients. We use the 90 GHz DMR map to provide the parameters of the prior for the 53 GHz sky map, and *vice versa*. We posit a Gaussian shape for the prior density of the $\ell \leq 2$ coefficients and assume independence among the components. We choose not to ignore the information about these lower-order terms that can be extracted from the data with a simple LS fit. However, since section 2 showed that the formal uncertainties of LS estimates are artificially small, we inflate the standard deviation of the $\ell \leq 1$ terms by a factor of two. Note that the dipole is 1000 times larger than the quadrupole and is less sensitive to Galactic cut effects (Lineweaver *et al* 1996). For the quadrupole terms, σ_2^2 and the lower bounds from table 2 are added in quadrature. For $\ell \geq 3$ we just use $\sigma_\ell^2(Q)$.

4.2. Power-law models with an RMS constraint

The conditional distribution of \mathbf{a} given \mathbf{d} is a function of the unknown Q . Before adopting prior distributions for the spectral parameters we study the sensitivity of the posterior estimates to a ‘reasonable’ range of Q values. The mild RMS uncertainty on the $a_{\ell m}$ was not strong enough to yield small quadrupole uncertainties. Is it enough if it is used to constrain Q ?

A priori plausible values for Q can be restricted using the observed RMS of the beam-convolved CMB: $\text{RMS}^2 = \sum_{\ell, m} G_\ell^2 a_{\ell m}^2 / 4\pi$. Under the model, the $\{a_{\ell m}\}$ are independent, and RMS^2 is approximately Gaussianly distributed. Let μ_{RMS^2} and $\sigma_{\text{RMS}^2}^2$ denote the mean and variance under the Gaussian approximation. The reported RMS estimate (Ba97) yields the following test: Reject the hypothesis that Q is the true value if the predicted RMS^2 for that Q is outside the approximate 95% confidence interval $\hat{\mu}_{\text{RMS}^2} \pm 1.96s_{\hat{\mu}_{\text{RMS}^2}}$. Figure 3 shows the resulting region R . For example, for $\eta = 1$ the interval for Q is consistent with the value $Q = 18 \pm 4 \mu\text{K}$ in Ba97.

We computed 68% credible regions for the quadrupole coefficients (appendix B) over a dense grid of values of Q in R . Taking R as a ‘reasonable’ constraint region for Q , the set of 68% Bayesian credible regions yields a conservative confidence region over R . Let a_{2m}^- and a_{2m}^+ be the smallest lower and larger upper endpoints, respectively, of any of the credible regions as Q varies over R . Define $I_m = [a_{2m}^-, a_{2m}^+]$, then for any $Q \in R$ we have

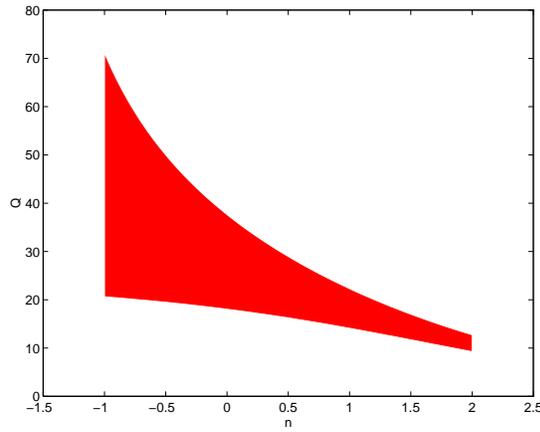


Figure 3. Region R for Q (μK) and η derived from a 95% confidence interval for the RMS.

Table 3. 68% confidence intervals for the quadrupole, and smallest and largest posterior uncertainties for $Q \in R$ (μK) for the 53 GHz and 90 GHz DMR maps. See section 4.2.

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_{RMS}
53	-18.5 ± 14.1 4.2, 7.3	38.0 ± 15.6 9.0, 11.6	11.5 ± 10.3 7.9, 9.5	1.8 ± 11.0 4.9, 9.6	8.7 ± 11.1 4.3, 8.2	12.6 ± 6.1 1.6, 3.4
90	23.4 ± 13.3 5.2, 7.2	17.4 ± 18.5 13.9, 16.2	28.0 ± 14.8 12.4, 13.6	-13.1 ± 16.3 6.5, 11.5	2.0 ± 10.2 5.9, 10.0	15.8 ± 7.2 2.4, 4.1

$P[I_m \ni \hat{a}_{2,m} | Q, \mathbf{d}] \geq 0.68$. Table 3 shows the I_m as well as the half-length of the shortest and longest 68% credible regions over R for each coefficient using the 20° e Galactic cut. The $\{a_{2m}\}$ have been normalized to Q_1, \dots, Q_5 . Note that the ratio of maximum to minimum confidence interval size, for $Q \in R$, could be as large as 2. This variability should be included in the posterior estimates. We do this next.

4.3. Assigning a prior to Q

One could argue that the uncertainties in table 3 are the best we can do given the information we have, but what we consider a ‘reasonable’ region is also subjective. To reduce the uncertainties in table 3 we need more information on Q , and this, given the lack of any other prior physical information, involves some kind of subjective choice. In this section we assume a joint prior distribution for the spectral parameters but this time we have no physical argument for the choice of prior. It is up to the reader to decide if the approach is a reasonable compromise between assuming only an RMS constraint, and assuming that they take fixed previously estimated values.

We posit two prior distributions on Q . In both cases, we take the prior distribution for the $\ell \leq 2$ components to be that described above, and use two different priors for Q .

Prior 1. Q and η are independent random variables with means and standard deviations equal to their DMR four-year estimates (section 2.1). An inverse gamma distribution is used for Q and a Gaussian for η . This choice of prior distributions is arbitrary, but a simulation study showed that the posterior estimates of quadrupole coefficients and Q are not sensitive to the shapes of the distributions provided the means and standard deviations are fixed.

Prior 2. To reduce the influence of the prior we also used the Jeffreys non-informative prior, $\propto \sqrt{\det \mathbf{I}(\mathcal{Q})}$, where \mathbf{I} is the Fisher information.

Once a prior distribution for \mathcal{Q} is chosen, the Bayes estimates of \mathbf{a} and Q_{RMS} , and their corresponding variances, are defined as the posterior quantities given the data, marginal on \mathcal{Q} . Gibbs sampling was used to compute the posterior estimates (appendix B.1). The final estimates using the 20°e Galactic cut are in table 4.

Table 4. Posterior quadrupole means \pm posterior standard deviations using the Gaussian and Jeffreys (j) priors (μK) on the 53 GHz and 90 GHz DMR maps.

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_{RMS}
53	-15.8 ± 7.3	38.5 ± 11.4	12.4 ± 9.3	0.0 ± 8.2	9.3 ± 7.5	11.6 ± 3.1
53 ^j	-16.1 ± 7.9	39.8 ± 11.5	12.4 ± 9.3	0.1 ± 8.3	9.5 ± 7.6	11.8 ± 3.4
90	22.3 ± 6.9	18.9 ± 15.8	27.7 ± 13.6	-8.7 ± 9.7	1.3 ± 8.5	14.1 ± 3.5
90 ^j	22.6 ± 7.2	18.7 ± 15.9	27.7 ± 13.6	-9.2 ± 10.2	1.3 ± 8.6	14.4 ± 3.8

4.4. Discussion

For each channel, the results for the two different priors are consistent. Estimates from the 53 GHz and 90 GHz maps are consistent at the 1σ level except for Q_1 , where they take opposite signs in the 90 GHz and 53 GHz maps. As pointed out by S92, this effect is an indication of residual Galactic contamination. Galactic emissions are stronger at 53 GHz than at 90 GHz.

K96 provide two tables of quadrupole results: table 3 corresponds to estimates for different frequency maps using no Galactic modelling, only the 20°e cut; table 4 reports the results for map combinations using Galactic cuts and Galactic templates. They provide two types of uncertainties: the ‘statistical’ uncertainties of the estimation procedure and the ‘systematic’ ones from the Galactic modelling. We compare ours to their statistical uncertainties. Figure 4 compares table 4 above to the cross-correlation results of table 4 in K96 and to the 90 GHz results from their table 3. Our quadrupole results are inconsistent with those in table 3 of K96, where, as in our case, no Galactic templates were used. A plausible explanation is that Galactic contamination in higher-order multipoles introduces a bias in their LS quadrupole estimates by adding a non-zero mean to their noise terms. This is consistent with the similarity of their Galactic template results with our table 4. However, their uncertainties for Q_2 and Q_3 seem too optimistic; this is inconsistent with the minimax results which assigned the largest uncertainty lower bounds to Q_2 and Q_3 . We have no explanation for this discrepancy but all of our estimates, including those in table 1, assign the largest uncertainties to those two coefficients.

5. Conclusions

Uncertainties in CMB quadrupole estimates depend crucially upon *a priori* information. Previous uncertainty estimates have either neglected power aliasing from $\ell \geq 3$ terms into the quadrupole or have tried to correct the aliasing using a power-law model with fixed parameters for the $\ell \geq 3$ spectrum. We have explored the dependence of quadrupole uncertainties upon different types of prior information, including: (i) *ad hoc* truncation of the spherical harmonic expansion; (ii) a constraint on the RMS fluctuations derived from observations; (iii) a two-parameter, power-law model for the spectrum of fluctuations with an RMS constraint on \mathcal{Q} ; (iv) a power-law model with a joint prior distribution on \mathcal{Q} . We have shown that: (i) LS

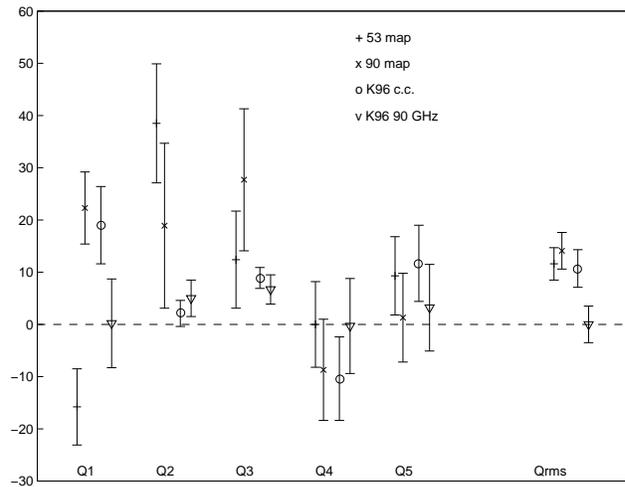


Figure 4. Estimated quadrupole parameters (μK) from table 4. Results from tables 3 and 4 in K96 are shown for comparison.

uncertainty estimates of individual quadrupole components are extremely sensitive to the degree of truncation; (ii) if CMB fluctuations are constrained only by the observed RMS and a 20° Galactic cut is used, lower bounds for 1σ detections of quadrupole coefficients are $|Q_1| \geq 5$, $|Q_2| \geq 71$, $|Q_3| \geq 69$, $|Q_4| \geq 52$, $|Q_5| \geq 50$ (μK). Even with the four-year maps, for the uncertainties to be as small as those in K96 on the basis of an RMS constraint alone would require at least a fivefold reduction in the noise per pixel; (iii) if CMB fluctuations are assumed to follow a power-law model with Q constrained by the observed RMS, the quadrupole uncertainties are about twice as large as those of K96; (iv) finally, only by assuming prior distributions on the spectral parameters are most of the uncertainties as small as those of K96. However, future data, for example those from the next generation satellites MAP and Planck, with much higher signal-to-noise ratio than that of DMR data will, in principle, lead to ‘small’ quadrupole uncertainties without the need for a cosmological model.

Acknowledgments

We are specially grateful to Costas Goutis, *in memoriam*, for many inspiring conversations. We are grateful to P J Bickel, S N Evans, A Gelman, J A Rice, J A Scales, J Silk and G Smoot for helpful conversations. PBS received support from NSF Presidential Young Investigator Award DMS-8957573 and grants DMS-9404276 and AST-9504410. LT was partially supported by grants DMS-9404276 and DGCYTPB94-0364. CHL acknowledges support from NSF/NATO 9552722.

Appendix A. Derivation and calculation of the minimax bounds

We re-scale the sky map data by the known σ_i and write (2) as $\mathbf{d} = \mathbf{K}\mathbf{a} + \mathbf{z}$, where \mathbf{a} is a sequence of spherical harmonic coefficients ordered to depend upon a single index, \mathbf{z} is multivariate normal $N(0, \mathbf{I})$ and \mathbf{K} is the data mapping, $(\mathbf{K}\mathbf{a})_i = \sum_l G_l a_l Y_l(\hat{r}_i)/\sigma_i$. The length of \mathbf{d} is the number of pixels outside the Galactic cut. We want to lower bound the length of $1 - \alpha$ confidence intervals of linear functionals $L\mathbf{a}$ (e.g. a quadrupole coefficient) of the

model \mathbf{a} , based on the data \mathbf{d} and the prior constraint $\|\mathbf{a}\|^2 \leq M^2$. If we base the confidence interval on the best possible estimator (the one that admits the shortest fixed-length confidence interval), its half-length is

$$C_\alpha = \inf\{\chi : \exists F \text{ s.t. } P(|F(\mathbf{d}) - L\mathbf{x}| \leq \chi) \geq 1 - \alpha, \|\mathbf{x}\| \leq M\}.$$

Donoho (1994) shows that

$$\frac{1}{2}w(2Z_{1-\alpha}) \leq C_\alpha \leq w(Z_{1-\alpha/2}) \tag{A.1}$$

where w is the modulus of continuity

$$w(\delta) = \sup\{\|L\mathbf{x} - L\mathbf{y}\| : \|K\mathbf{x} - K\mathbf{y}\| \leq \delta, \|\mathbf{x}\|, \|\mathbf{y}\| \leq M\}.$$

An estimate of the bound M in (3) is obtained from the RMS CMB estimate in Ba97: $\int T^2(\hat{r}) d\Omega \cong 4\pi \text{RMS}^2 \equiv M^2$. The modulus of continuity for estimating the k th harmonic coefficient with the given quadratic bound is then

$$w_k(\delta) = -\inf\{-x_k : \|K\mathbf{x}\|^2 \leq \delta^2, \|\mathbf{x}\|^2 \leq 4M^2\}. \tag{A.2}$$

We discretized (A.2) using a sequence of finite-dimensional moduli of continuity $w_k(\ell_{\max}, \delta)$, defined by retaining only those terms of degree $\leq \ell_{\max}$ in (A.2). By construction, $w_k(\ell_{\max}, \delta)$ is a non-decreasing sequence in ℓ_{\max} converging to $w_k(\delta)$. In the discrete approximation, (A.2) becomes

$$w_k(\ell_{\max}, \delta) = -\inf\left\{-\mathbf{e}_k^t V\mathbf{y} : \sum_i^{N(\ell_{\max})} \lambda_i^2 y_i^2 \leq \delta^2, \sum_i^{N(\ell_{\max})} y_i^2 \leq 4M^2\right\}$$

where \mathbf{e}_k is a vector of zeros, except for a 1 in the k th entry, the columns of V form an orthonormal basis of eigenvectors of $K^t K$ with eigenvalues (λ_i) , and $\mathbf{x} = V\mathbf{y}$. A Householder transformation was used to reduce $K^t K$ to tridiagonal form. The QL algorithm was then used to find V and $\{\lambda_i\}$. The finite-dimensional optimization problems were solved with the Stanford systems optimization code NPSOL (Gill and Murray 1986). Each chosen ℓ_{\max} yields a lower bound for the length of the confidence interval. Although the constraints are diagonal, NPSOL uses a dense matrix to approximate the reduced Hessian. This limited the maximum practical ℓ_{\max} for which we could solve the problem numerically. Our strategy was to wait for the solution to be on the boundary of the constraining ellipsoids and then to increase ℓ_{\max} until the solution stabilized.

Appendix B. Bayesian calculations

Consider an approximation of degree ℓ_{\max} to (2): $\mathbf{d}_{\ell_{\max}} \sim \mathbf{X}\mathbf{a} + \mathbf{z}$, where \mathbf{X} uses only those terms of degree $\leq \ell_{\max}$. The choice of ℓ_{\max} is not important since the posterior conditional mean of \mathbf{a} stabilizes for a sufficiently large ℓ_{\max} (figure B1). The conditional density of $\mathbf{d}_{\ell_{\max}}$ given \mathbf{a} is a multivariate normal $P(\mathbf{d}_{\ell_{\max}}|\mathbf{a}, \mathcal{Q}) \sim N(\mathbf{X}\mathbf{a}, \mathbf{I})$. Suppose that \mathcal{Q} is known. Within the finite-dimensional approximation, the Bayes estimate of $a_{\ell m}$ is the posterior mean of $a_{\ell m}$ given $\mathbf{d}_{\ell_{\max}}$ and \mathcal{Q}

$$\hat{\mathbf{a}} = E(\mathbf{a}|\mathbf{d}_{\ell_{\max}}, \mathcal{Q}) = (\mathbf{X}^t \mathbf{X} + \Sigma_a^{-1})^{-1} (\mathbf{X}^t \mathbf{d}_{\ell_{\max}} + \Sigma_a^{-1} \boldsymbol{\alpha}) \tag{B.1}$$

$$\text{Cov}(\hat{\mathbf{a}}) = \text{Cov}(\mathbf{a}|\mathbf{d}_{\ell_{\max}}, \mathcal{Q}) = (\mathbf{X}^t \mathbf{X} + \Sigma_a^{-1})^{-1} \tag{B.2}$$

where Σ_a is the prior covariance matrix of \mathbf{a} . The posterior mean and standard deviation of Q_{RMS} are computed through simulations using the posterior distribution of \mathbf{a} . Note that (B.1) and (B.2) are like ridge regression with a weighted spectrum, $\mathbf{a}^t \Sigma_a \mathbf{a}$, as a penalty; estimates are thus expected to stabilize for large ℓ_{\max} . Figure B1 shows posterior estimates of Q_{RMS} as a function of ℓ_{\max} for the 53 GHz map with a 20° Galactic cut and prior model $\mathcal{Q} = (15.3, 1.2)$. The estimates are stable for $\ell_{\max} \geq 10$.

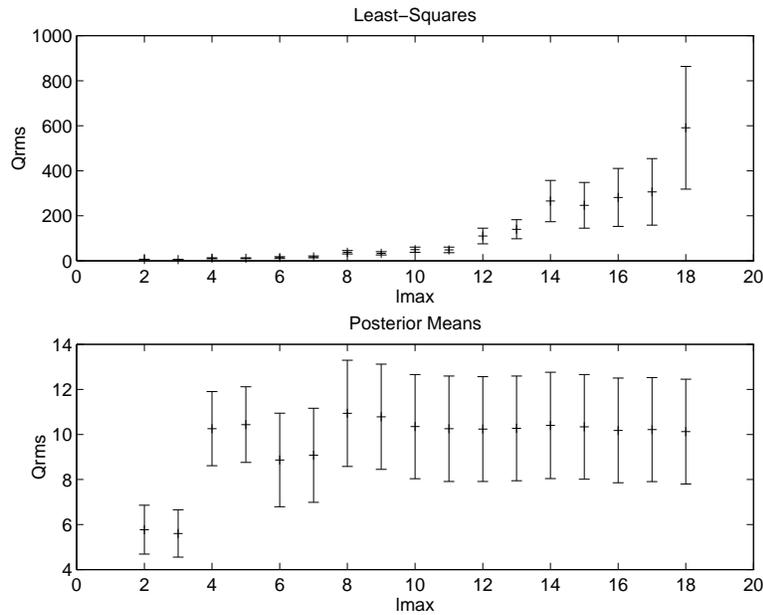


Figure B1. LS (top) and Bayesian (bottom) estimates of Q_{RMS} (μK) as a function of ℓ_{max} (20° Galactic cut). Bayesian estimates are obtained using (B.1) and (B.2) to estimate the posterior distribution of Q_{RMS} .

B.1. Computing the Bayes estimators by Markov chain Monte Carlo methods

Gibbs sampling was used to draw samples from $P(\mathbf{a}|\mathbf{T})$ indirectly, using the conditional distributions. Starting with initial values for \mathcal{Q} , we draw \mathbf{a} from $P(\mathbf{a}|\mathbf{T}, \mathcal{Q})$. The resulting \mathbf{a} is then used to draw \mathcal{Q} from $P(\mathcal{Q}|\mathbf{a}, \mathbf{T}, \eta)$. Finally, η is drawn from $P(\eta|\mathbf{a}, \mathbf{T}, \mathcal{Q})$, and the process starts again. Ultimately, we obtain samples of \mathbf{a} , \mathcal{Q} and η from their corresponding posterior marginal distributions conditional on \mathbf{T} . Since no closed expressions for $P(\mathcal{Q}|\mathbf{T}, \mathbf{a}, \eta)$ and $P(\eta|\mathbf{T}, \mathbf{a}, \mathcal{Q})$ are available, we used a Metropolis algorithm with symmetric jumping densities and acceptance rates of about 40% (Gelman *et al* 1996). One further difficulty is a ridge in \mathcal{Q} space that forces the Gibbs sampler to meander inside a ‘half-moon’ shape. To improve convergence, we renormalized the harmonics to the k -multipole for which the information matrix was closer to a diagonal. Convergence was checked by monitoring cumulative sample means and by changing the starting values of the chain.

References

- Banday A *et al* 1997 Root mean square anisotropy in the COBE DMR four-year sky maps *Astrophys. J.* **475** 393
 Bennett C *et al* 1996 4 Year COBE DMR cosmic microwave background observations: maps and basic results *Astrophys. J.* **464** L1
 Bond J and Efstathiou G 1987 The statistics of cosmic background fluctuations *Mon. Not. R. Astron. Soc.* **226** 655–87
 Bunn E, Hoffman Y and Silk J 1994 The effect of incomplete sky coverage on the analysis of large angular scale microwave background anisotropy *Astrophys. J.* **425** 359
 Donoho D 1994 Statistical estimation and optimal recovery *Ann. Stat.* **22** 238–70
 Gelman A, Roberts G and Gilks W 1996 *Bayesian Statistics* vol 5 (Oxford: Oxford University Press) p 599
 Gill P and Murray W 1986 *Technical Report* 86-2 Stanford University, Department of Operations Research
 Gould A 1993 An estimate of the COBE quadrupole *Astrophys. J.* **403** L51

- Kogut A *et al* 1996 Microwave emission at high Galactic latitudes in the four-year DMR sky maps *Astrophys. J.* **464** L5
- Lineveaver C *et al* 1994 Correlated noise in the COBE DMR sky maps *Astrophys. J.* **436** 452
- 1996 The dipole observed in the COBE DMR four-year data *Astrophys. J.* **470** 38
- Pulliam R and Stark P B 1993 Bumps on the coremantle boundary: are they facts or artifacts? *J. Geophys. Res.* **98** 1943–56
- 1994 Confidence regions for mantle heterogeneity *J. Geophys. Res.* **99** 6931–43
- Silk J 1997 *A Short History of the Universe* (New York: Freeman)
- Smoot G *et al* 1992 Structure in the COBE DMR first year maps *Astrophys. J.* **396** L1–5
- Stark P B 1992 Minimax confidence intervals in geomagnetism *Geophys. J. Int.* **108** 329–38
- 1993 Uncertainty of the COBE quadrupole detection *Astrophys. J.* **408** L73
- 1995 Reply to a comment by Morelli and Dziewonsky *J. Geophys. Res.* **100** 15
- Stark P B and Hengartner N 1993 Reproducing Earth's kernel: uncertainty of the shape of the core–mantle boundary from PKP and Pcp travel times *J. Geophys. Res.* **98** 1957–72
- Wright E *et al* 1994 Comments on the statistical analysis of excess variance in the COBE differential microwave radiometer maps *Astrophys. J.* **420** 1–8