Working Paper 95-32
Statistics and Econometrics Series 09
September 1995

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

ESTIMATING PARAMETERS OF FLUCTUATIONS IN THE COSMIC MICROWAVE BACKGROUND

Luis Tenorio, Charles H. Lineweaver and George Smoot*

Abstract

We address questions that arise in statistical analyses of recently detected fluctuations in the Cosmic Microwave Background (CMB). Estimators of the quadrupole amplitude, Q, and spectral index, n, of the CMB angular fluctuation power spectrum are considered. Families of unbiased estimators of Q^2 and existence conditions for minimum variance estimators of n are given. We find that the common practice of excluding the quadrupole is not recommended if one is interested in unbiased estimators. We explain previously reported correlations of the estimators and show how they depend on the multiple used to normalize the spectrum. We show that a finite beam resolution does not justify the use of truncated least-squares to estimate harmonic coefficients of CMB data.

Key Words

COBE; Data analysis; Unbiased Estimators; Maximum likelihood; Monte Carlo

*Tenorio, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid y Space Sciences Laboratory, Berkeley, CA, USA; Lineweaver, Observatoire Astronomique de Strasbourg, France; Smoot, University of California, Berkeley, CA, USA.

1. Introduction

Thirty years ago the Cosmic Microwave Background (CMB) radiation was discovered by Penzias and Wilson (1965). This nearly isotropic, 2.7 K blackbody radiation is the redshifted relic of the Big Bang, and is comprised of the oldest observable photons. Their long journey towards us has lasted 99.997% of the age of the universe; a journey that began when the photons were last scattered by free electrons of the cosmic plasma, at a time when the universe was 1000 times smaller and the temperature 1000 times higher than the CMB today. Observed large scale structures like galaxies, clusters of galaxies, giant voids and superclusters are believed to be the result of gravitational instabilities of under- and over-densities in the matter distribution. If that is the case their imprint must be present in the form of anisotropies in the temperature distribution of the CMB. For 25 years, astrophysicists had been looking for these anisotropies, which are crucial ingredients in present theories of large scale structure formation. In 1989, NASA launched its first cosmology satellite, COBE. The Differential Microwave Radiometer (DMR) instrument on board COBE was constructed to search for these 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.

Since our observable universe is a single finite sample of a larger universe, models of large-scale structure can only make statistical predictions about the observable CMB fluctuations. This statistical uncertainty in the predictions is usually known as cosmic variance. For example, a popular class of models predicts Gaussian distributed spherical harmonic coefficients of the fluctuations. More precisely, CMB fluctuations in a given direction in the sky are written in terms of real spherical harmonics as (Smoot *et al.* 1991)

$$\Delta T/T(\theta,\phi) = \sum_{l,m} a_{lm} F_{lm}(\theta,\phi), \qquad (1)$$

where the coefficients a_{lm} are modeled as idependent $N(0,\sigma_l^2)$ variables, with σ_l being a function of two parameters; the spectral index n of the power spectrum of the matter density fluctuations, and its amplitude Q. Usually Q is written as Q_{rms-PS} to emphasize that it is the rms amplitude of the quadrupole ($l = 2$) normalized power spectrum. If n can be determined to be greater than one, then some favorite theories (e.g. some inflationary models) will have to be abandoned or modified. On the other hand, if n is found to be less than one, then the contribution of gravitational waves towards the production of fluctuations could be significant. There have been a number of studies to estimate the spectral parameters that best fit the data. Figure 1 is a summary of published n and Q estimates from COBE DMR data. Many of the estimated values of Q in Fig. 1 have been obtained by simply fixing $n = 1$. Also, the best fitting quadrupole is sometimes removed from the data before any estimation is done (e.g. Lineweaver *et al.* 1994, Smoot *et al.* 1994, Wright *et al.* 1994). This quadrupole exclusion is motivated by the possibility of galactic contamination and systematic error contributions to the measured quadrupole. Exclusion of the quadrupole results in values of n that are closer to the favorite inflationary value $n = 1$. When the quadrupole is included, larger n values are obtained as the observed quadrupole appears significantly lower than that predicted by fluctuation amplitudes at higher orders; only the five points to the right in Fig. 1 include the quadrupole. The curves in Fig. 1 are fits to scatter plots of Monte-Carlo estimates of the spectral parameters. This correlation of the estimators makes it difficult to estimate n and Q with Monte-Carlo methods.

This work was motivated by questions arising in the analysis of DMR data: Do there exist unbiased estimators of Q^2 besides the observed total power at $l = 2$? Do minimum variance unbiased estimators of n and Q exist? How good are maximum likelihood (ML) methods for the CMB problem, consisting of non-identically distributed observations? How are estimates affected by the exclusion of the quadrupole? Why are reported estimators always correlated? Are least-squares estimates of harmonic coefficients affected by the order of the fit? To answer these questions we consider estimators that can be expressed as functions of finitely many spherical harmonic coefficients. For example, Q^2 can be estimated using the total power, $\sum a_{2m}^2$, at $l = 2$ or using ML of the a_{lm} (e.g. Gorski *et al.* 1994). These estimators also arise naturally when Monte-Carlo methods are used: given a map of temperatures $T_1, ..., T_k$, a statistic $S(T_1, ..., T_k)$ is defined (Banday *et al.* 1994, Kogut *et al.* 1995, Smoot *et al.* 1994). Then its mean $\mu_S(n, Q)$ and variance-covariance $\Sigma_S(n, Q)$ are determined through sky map simulations for different n and Q incorporating actual noise levels, observation patterns, known systematics and pixelization. Next, a criterion is defined for choosing the best n and Q by comparing the measured S to the one obtained from simulations. Usually estimators $\hat{n}$ and $\hat{Q}$ are defined as those values that minimize or maximize a test statistic (e.g. χ^2), which is formed using the simulated μ_S and Σ_S, and the statistic S of the real data. However, the statistic S of the observed data depends in principle on infinitely many a_{lm} while μ_S and Σ_S are computed using only those of order $l \leq L$. We may assume that the estimators are functions of a_{lm}, $l \leq L$, provided they are insensitive to this order truncation.

Among those statistics given in terms of finitely many a_{lm}, unbiased and ML estimators are considered. The a_{lm} will be assumed to be noiseless, subject only to cosmic variance, providing a reference as to the best one might hope to achieve using these estimators. Establishing the existence of uniformly minimum unbiased estimators (UMVU) estimators, or even of unbiased estimators, for a given maximum order L is not straightforward. Some existence conditions are given in Section 2. On the other hand, ML estimators of n and Q^2 are relatively easy to compute numerically. The use of ML is usually justified by its nice asymptotic properties. However, one has to be cautious since these properties are usually proved under the assumptions of independent and identically distributed variables. The a_{lm} are not identically distributed for different l. Although there are some results concerning the non-identically distributed case (e.g. Hoadley 1971), we conducted Monte-Carlo simulations to study ML estimators of n and Q^2 for the values of L frequently used, and to check for possible systematic effects caused by the exclusion of the quadrupole. Section 3 presents results of our simulations, as well as a discussion of the observed bias, variance, and correlations.

In the last section we point out some difficulties in estimating the a_{lm} themselves. To estimate harmonic coefficients of our observable universe, cosmic variance is neglected and only variance contributions due to the measurement process are considered. Suppose we want to estimate the monopole and dipole ($l = 0, 1$) coefficients. This is usually done by least-squares fitting the signal to a harmonic expansion of some finite order $L \geq 1$. It is sometimes believed that the finite resolution of the beam allows one to safely assume a finite harmonic expansion of the data, i.e. the a_{lm} estimates are not order dependent if L is high enough. Harmonic coefficients of a linear combination of two functions on the sphere are linear combinations of the corresponding coefficients of each function, i.e., the a_{lm} are linear functionals of the temperature. It has been shown that only linear functionals which are linear combinations of the data generating process can be recovered with finite uncertainty in a linear inverse problem, unless additional constraints are imposed (Backus 1970). Due to galactic plane cuts, uneven sky sampling and beam resolution, harmonic coefficients of DMR data do not have this property. Uncertainties of the a_{lm} estimators depend on L. This truncation effect has been pointed out before (e.g. Backus 1988, Benton *et al.* 1982, Stark 1993) but a simple toy example is illustrative. The last section gives a simple COBE-like example to show how misleading least-squares estimates of the a_{lm} can be.

2. The Model

Power-law models predict a linear power spectrum of the CMB fluctuations for the angular scales probed by the DMR instrument ($\gtrsim 7°$). Such a spectrum is usually parametrized by an amplitude Q and a spectral index n. Assume that harmonic coefficients a_{lm} in the expansion (1) are noiseless, subject only to cosmic variance. Then the a_{lm} are modeled as independent $N(0, \sigma_l^2)$ variables, with cosmic variance σ_l^2 given by

$$\sigma_l^2 = Q^2 N_l(n), \quad N_l(n) = \frac{1}{5}\frac{\Gamma(l+(n-1)/2)\Gamma((9-n)/2)}{\Gamma(l+(5-n)/2)\Gamma((3+n)/2)}, \tag{2}$$

for $l > 1$, $Q > 0$. The functions N_l have a singularity at $n = -3$ and are only physically meaningful for $n < 3$ (Efstathiou 1990, Lineweaver 1994). We therefore consider the range $-3 < n < 3$ as a natural parameter space for n, but the more restricted interval $-1 \leq n \leq 2$ is physically reasonable in that it spans all the experimentally determined values of n.

The joint distribution of the a_{lm} for $l \leq L$ can be written as

$$f(a_{2,-2}, ..., a_{L,L}) \propto \frac{1}{\prod_i Q\sqrt{N_i(n)}} \exp[-\frac{1}{2}\sum_l \frac{(2l+1)P_l}{Q^2 N_l(n)}], \tag{3}$$

where

$$P_l = \frac{1}{2l+1}\sum_{m=-l}^{l} a_{lm}^2 \sim \frac{Q^2 N_l(n)}{2l+1}\chi^2_{2l+1}.$$

It follows that the spectrum, $\mathbf{P} = (P_2, ..., P_L)$, is sufficient for Q and n.

3. Unbiased Estimation

Consider unbiased estimators of the spectral parameters that are functions of the observed a_{lm} ($l \leq L$). Since $\mathbf{P}$ is sufficient, unbiased estimators of n and Q^2 may be assumed to be functions of $\mathbf{P}$. We will require that estimators $\hat{n}$ and $\hat{Q}^2$ behave properly under re-scalings, i.e., multiplying all the a_{lm} by a constant c should scale the estimated Q^2 with a c^2 while the estimated n should remain unchanged. This means that for any $k > 0$

$$\hat{Q}^2(k\mathbf{P}) = k\hat{Q}^2(\mathbf{P}) \quad \text{and} \quad \hat{n}(k\mathbf{P}) = \hat{n}(\mathbf{P}). \tag{4}$$

Since $N_2(n) = 1/5$, we have that $\hat{Q}_o^2 = 5P_2$ is an unbiased estimator of Q^2. But $\hat{Q}_o^2$ has a large variance and it is rarely used to estimate Q^2. However, we know of no other unbiased estimator of Q^2 previously reported. It is natural to ask whether $\hat{Q}_0^2$ is the only one. Take the simple case $L = 3$, i.e., $\mathbf{P} = (P_2, P_3)$. By completeness, $\hat{Q}_0^2$ is the only unbiased estimator of Q^2 and therefore it is also UMVU. A larger number of a_{lm} could be used by choosing P_2 and any other P_l instead, but $\hat{Q}_0^2$ will still be the unique UMVU estimator of Q^2. It is shown in the Appendix (Pf.1) that these are the only cases for which $\hat{Q}_0^2$ is an UMVU estimator. Any unbiased estimators of Q^2 can be written as $\hat{Q}^2 = \hat{Q}_o^2 - U(\mathbf{P})$, for some unbiased estimator of zero $U(P)$. That there do exist families of unbiased estimators of zero for $L > 3$ is shown in the Appendix (Pf.1). We thus have infinitely many unbiased estimators of Q^2 when $L > 3$. The question is whether there exists one with uniformly minimum variance. So far we are only able to answer that $\hat{Q}_0^2$ is not UMVU if $L > 3$ (Appendix Pf.1). In fact, no P_l is an UMVU estimator of its expectation. But, the question of whether there exist UMVU estimators of Q^2 for $L > 3$ is still unsolved.

Whether an unbiased estimator of the spectral index n exists is still unknown. But, given L, we ask how many of the P_l ($l \leq L$) are required to form an UMVU estimator. Such an estimator of n, if there is one, can not be independent of two consecutive P_l, P_{l+1} (Appendix Pf.2). Therefore, if an UMVU estimator of n exists for $l \leq L$, it will not be UMVU if L is increased to $L + 2$. It is also clear that such an estimator of n will depend on at least half of the P_l, $l \leq L$.

As mentioned in the Introduction, the quadrupole is sometimes removed from the data before estimating n and Q. How are unbiased estimatators of the spectral parameters affected by the exclusion of P_2? Heuristically, as $n \to -3$ the expected amplitudes of each of the P_l goes to zero, except that of P_2. Thus, if P_2 is not included we are only left with the vanishing P_l. Not surprisingly, there does not exist an unbiased estimator of Q^2 for the parameter space $Q > 0$, $-3 < n < 3$ (Appendix Pf.3). However, it is still not clear whether unbiased estimators of Q^2 exist for subintervals of $-3 < n < 3$ when P_2 is excluded.

4. Maximum Likelihood Estimation

Solving the maximum likelihood equations leads to estimators $\hat{n}_{\rm ml}$ and $\hat{Q}^2_{\rm ml}$ that are functions of the sufficient statistic $\mathbf{P}$ and that satisfy the scaling conditions (4). For example, when $L = 3$ the ML estimators are

$$\hat{Q}^2_{\rm ml} = 5P_2, \quad \hat{n}_{\rm ml} = \frac{9P_3 - 3P_2}{P_3 + P_2}.$$

In this case, $\hat{Q}^2_{\rm ml}$ is the unbiased estimator considered in Section 3, but $\hat{n}_{\rm ml}$ is biased: a short analytical calculation shows that $E(\hat{n}_{\rm ml}) = 0.489$ and 1.409 for $n = 0$ and 1, respectively. It is also easy to see that $\hat{n}_{\rm ml}$ and $\hat{Q}^2_{\rm ml}$ are negatively correlated, in fact,

$$E(\hat{Q}^2_{\rm ml}|\hat{n}_{\rm ml}) = 12Q^2 \frac{N_3(n)}{5N_3(n) + 7N_3(\hat{n}_{\rm ml})}.$$

Figure 2 shows a scatter plot of $\hat{n}_{\rm ml}$ and $\hat{Q}^2_{\rm ml}$ for simulated P_l with input values $n_0 = 1$, $Q_0 = 1$. The solid curve is the conditional expectation of $\hat{Q}^2_{\rm ml}$ given $\hat{n}_{\rm ml}$. The curves in Fig. 1 are basically estimates of the corresponding conditional expectation.

4.1. Simulation Results and Discussion

We study ML estimators for higher values of L through Monte-Carlo simulations. The ML estimator of n is determined by finding the zero of $\partial_n log(\mathrm{L}[n, Q])$, where $\mathrm{L}[n, Q]$ is the likelihood obtained from Eqn.(3). We use a combination of the Newton-Raphson and bisection methods to determine the zero to an accuracy of five decimal places. Once $\hat{n}_{\rm ml}$ is found, the ML estimator of Q^2 is determined by substitution into the Q-derivative of the log-likelihood equation. Since the estimators satisfy the scaling conditions, it suffices to consider the case $Q = 1$. Due to the resolution of the DMR beam, Monte-Carlo simulations of DMR data are usually performed using harmonic expansions to $L \sim 30$. The fits in the simulations here are for $L = 20$ and $L = 30$.

Figures 3a-d show estimated biases of $\hat{n}_{\rm ml}$ and $\hat{Q}_{\rm ml}$ as functions of n. Both estimators show a positive bias. The n dependence of the bias of Q seems to be large while the n bias is almost n independent. For $L = 30$ the average biases are about 0.004 and $0.001Q$ for $\hat{n}_{\rm ml}$ and $\hat{Q}_{\rm ml}$ respectively. Excluding the quadrupole (solid symbols) does not seem to affect the bias of $\hat{n}_{\rm ml}$ but increases that of $\hat{Q}_{\rm ml}$.

We now compare the variance of the estimators with their Cramèr-Rao (CR) bound. The information matrix I of Q^2 and n is

$$I(Q^2, n) = \begin{pmatrix} \frac{1}{2Q^4} \sum_{l \le L}(2l+1) & \frac{1}{2Q^2} \sum_{l \le L}(2l+1)\frac{\partial}{\partial n} ln(N_l) \\ \frac{1}{2Q^2} \sum_{l \le L}(2l+1)\frac{\partial}{\partial n} ln(N_l) & \frac{1}{2} \sum_{l \le L}(2l+1)(\frac{\partial}{\partial n} ln(N_l))^2 \end{pmatrix}. \quad (5)$$

Therefore, the CR bounds for n, Q and Q^2 are

$$\mathrm{CR}(n) = I_{11}/\det(I), \quad \mathrm{CR}(Q^2) = I_{22}/\det(I), \quad \mathrm{CR}(Q) = \mathrm{CR}(Q^2)/4Q^2.$$

Figures 4a and 4b compare the estimated standard errors of $\hat{n}_{\mathrm{ml}}$ and $\hat{Q}_{\mathrm{ml}}$ with the square root of their CR bound. The plots show that the ML estimators are basically efficient for $L = 20$ and $L = 30$. The distance from the bound does not seem to be affected by the exclusion of the quadrupole. We show next that this efficiency is partly responsible for the observed correlation between n and Q.

4.2. Correlation of the Estimators

The estimators $\hat{n}_{\mathrm{ml}}$ and $\hat{Q}^2_{\mathrm{ml}}$ are negatively correlated when $L = 3$. Simulations show that this is also the case for higher values of L. All reported estimators of n and Q have also shown this correlation (e.g. curves in Fig. 1), which makes difficult the determination of n and Q if a Monte-Carlo method like the one described in the Introduction is used. Some studies have reported a fit to the curve $E(\hat{Q}|\hat{n})$ (see Fig. 1) and some others have claimed that the correlation is due to instrument noise in the data (Smoot *et al.* 1994). But, as Fig. 1 shows, the correlation is present in ML estimators of n and Q^2 even for noiseless data. The observed correlation is a property of the estimators. Independent estimators of the spectral parameters do exist. For example, by independence of the P_l, independent estimators with small bias could be found by choosing a high enough L and determining the ML estimator of n using only the odd P_l, and the ML estimator of Q^2 using only the even ones. The CR bound can be used to find a simple sufficient condition to explain the observed negative correlation. Consider the CR bound of $n - Q^2$: $\mathrm{CR}(n - Q^2) = (I_{22} + I_{11} + 2I_{12})/\det(I)$, where I is the information matrix in Eqn.(5). It is easy to see that $I_{12} > 0$ for $-3 < n < 3$. Therefore,

$$\mathrm{CR}(n - Q^2) - \mathrm{CR}(n) - \mathrm{CR}(Q^2) = 2I_{12}/\det(I) \geq 0. \tag{6}$$

Now, take any estimators $\hat{n}$ and $\hat{Q}^2$ such that their variances and the variance of $\hat{n} - \hat{Q}^2$ achieve their corresponding CR bounds. It then follows from Eqn.(6) that

$$\mathrm{var}(\hat{n} - \hat{Q}^2) - \mathrm{var}(n) - \mathrm{var}(Q^2) = -2\mathrm{cov}(\hat{n}, \hat{Q}^2) = 2I_{12}/\det(I) \geq 0.$$

Therefore $\mathrm{cov}(\hat{n}, \hat{Q}^2) \leq 0$, and $\hat{n}$ and $\hat{Q}^2$ are negatively correlated. The same is true if the variances are not equal but 'close enough' to their CR bounds. This is precisely the case of the ML estimators found above. In other words, any estimators of the two spectral parameters which are based on the same statistic **P** and which are almost efficient will be negatively correlated. A similar argument predicts the sign of the correlation for different power spectrum normalizations. Suppose we normalize the spectrum to the kth multipole. Take R to be the new normalizing amplitude, and $E(P_l) = R^2 M_l(n)$. It can be shown (Appendix Pf.4) that almost efficient estimators of R and n are negatively or positively correlated depending on whether $\partial_n ln(N_k)$ is smaller or larger than the weighted l-average of $\partial_n ln(N_l)$. E.g., it will be negative for small values of k and positive for larger ones. One can use Eqn.(6) to find a k that gives a small correlation in order to get a better Monte-Carlo estimate of n.

5. Covariance of Harmonic Coefficients

COBE measurements of the CMB are convolutions of the DMR instrument's beam pattern with a sky signal $\Delta T/T$. The relationship between the true sky spherical harmonic coefficients a_{lm} of $\Delta T/T$ and those of the convolved sky b_{lm} is (ignoring pixelization effects)

$$b_{lm} = G_l a_{lm},$$

where the G_l are known Legendre coefficients of the beam pattern. Reported estimates of the a_{lm} are usually G_l-corrected least-squares fits of the b_{lm}. It is usually believed that the expansion of the real signal can be assumed finite due to the exponential decay of the G_l. However, using only a finite number of harmonics in the least-squares fits effectively adds the extra condition of having a finite harmonic expansion. Using least-squares to estimate the coefficients leads to estimators that depend on this extra information. For example, we fit a dipole to a 2-year 53 GHz COBE sky map using least-squares to different orders and galactic cuts. Figure 5 shows the estimated x-component amplitude of the dipole in celestial coordinates and its estimated error. With a 5° galactic cut, the absolute value of the x component amplitude decreases with L, while its estimated error increases. Due to larger correlations of the harmonics on the further cut sphere, the increase in the error is steeper with 10° galactic cut fits. Instabilities in the matrix operations may have something to do with the observed effects, especially for larger values of L. A simple example in which all the matrix operations can be done by hand may give more convincing evidence of a truncation order dependence.

5.1. *A Toy Example*

Start with a non-orthonormal basis $\{\chi_l(t)\}$, $0 \le t \le 1$ (non-orthogonality of the harmonics for DMR data is mainly due to galactic plane cuts and uneven sky sampling)

$$\chi_1(t) = e_1(t),\ \chi_l(t) = e_l(t) - e_{l-1}(t),\ l > 1,$$

where $\{e_i(t)\}$ is any orthonormal basis. Suppose we have measurements $y(t)$ modeled by

$$y(t_i) = \sum_{l=1}^{L} a_l G_l\, \chi_l(t_i) + \sum_{l=L+1}^{\infty} a_l G_l\, \chi_l(t_i) + \epsilon(t_i),$$

where $\epsilon(t)$ are iid $N(0,\sigma^2)$ and $t_i = i/I$, $1 \le i \le I$, index the different observations. Assume that the G_l are so small for $l > L$ that only the first L terms are considered in the fit. We show that no matter how small the G_l may be, truncating the expansion to a finite order leads to estimates of a_l that strongly depend on L.

A least-squares fit to tha data using only bases functions up to order L is obtained by minimizing $||\mathbf{y} - \mathbf{X}\mathbf{b}||^2$, where $b_l = a_l G_l$ and $X_{li} = \chi_l(t_i)$ for $1 \le l \le L$ and $1 \le i \le I$. The least-squares estimator of b is

$$\hat{\mathbf{b}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{y}.$$

Assume that sampling in t is fine and uniform enough to assure that sums of products over the I observations are well approximated by the scalar product of the corresponding vectors in t-space (e.g. $\sum_{i=1}^{I} \chi_l(t_i)\chi_m(t_i) = I < \chi_l, \chi_m >$). The matrix $(\mathbf{X}^t\mathbf{X})^{-1}$ can then be written explicitly

$$(\mathbf{X}^t\mathbf{X})^{-1} = \frac{1}{I}\begin{pmatrix} L & L-1 & L-2 & L-3 & . & . & . & 1 \\ L-1 & L-1 & L-2 & L-3 & . & . & . & 1 \\ L-2 & L-2 & L-2 & L-3 & . & . & . & 1 \\ . & & & & & & & . \\ . & & & & & & & . \\ . & & & & & & & . \\ 1 & 1 & 1 & 1 & . & . & . & 1 \end{pmatrix}.$$

Notice that the entries of $(\mathbf{X}^t\mathbf{X})^{-1}$ grow linearly with L. The estimator $\hat{\mathbf{b}}$ is biased

$$\text{bias}(\hat{\mathbf{b}}) = E(\hat{\mathbf{b}} - \mathbf{b}) = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{G},$$

where

$$\mathbf{G}^t = I(\sum_{l>L} a_l G_l < \chi_1, \chi_l >, ..., \sum_{l>L} a_l G_l < \chi_L, \chi_l >).$$

In particular,

$$|\text{bias}(\hat{b}_l)| = |a_{L+1}| G_{L+1} \quad \text{and} \quad |\text{bias}(\hat{a}_l)| = |a_{L+1}| \frac{G_{L+1}}{G_l}.$$

For example, the bias is small for large L if the a_l are known to be uniformly bounded, i.e., increasing the order decreases the bias. The covariance is not so well behaved; the covariance matrices of $\hat{b}$ and $\hat{a}$ are

$$\Sigma_{\hat{b}\hat{b}} = \sigma^2(\mathbf{X}^t\mathbf{X})^{-1} \quad \text{and} \quad \Sigma_{\hat{a}\hat{a}} = \sigma^2 \mathbf{D}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{D},$$

where $\mathbf{D} = \text{diag}\{G_1^{-1}, ..., G_L^{-1}\}$. It is clear that the variance should grow as $1/G_l$ because the G_l are decaying, making higher order terms weakly determined; and that it should decrease as the resolution I increases. But, for a fixed l, the variance and covariance also grow linearly with L. For example, since $\text{var}(\hat{a}_l) = (L+1-l)/(G_l I)$, the variance ratio of $\hat{a}_1$ for fits of order $l = L$ and $l = 1$ is just L. This does not mean, however, that it is best to do lower order fits. Then we really get artificial uncertainties to an ill determined inverse problem. Work to overcome these difficulties with the COBE data is still in progress. A possible solution might be to add constraints from prior physical information. For example, quadratic bounds have been used for similar problems in geomagnetism (e.g. Backus 1989, Stark 1992).

6. Conclusions

We have chosen a class of estimators which arises naturally in the type of Monte-Carlo methods used in COBE DMR data analyses. Within this class we considered unbiased and ML estimators. We found families of unbiased estimators of Q^2 but were unable to prove the existence of UMVU estimators. No unbiased estimators of n were found but existence conditions for UMVU estimators were given. It was shown that excluding the quadrupole from the data introduces a bias in estimators of Q^2, but does not affect the efficiency of ML estimators. We showed that the observed negative correlation of reported estimates of n and Q is a property of the estimators used; any two almost efficient estimators of the spectral parameters that are based on the same sufficient statistic $\mathbf{P}$ are negatively correlated. In particular, Monte-Carlo estimates of n and Q which are insensitive to truncation order and which are almost efficient will be correlated, depending on the multipole to which the power spectrum is normalized. It was also shown that despite the finite beam resolution, least-squares fits of harmonic coefficients of DMR data should be corrected for an order dependence of the fit.

Aknowledgements

We greatfully acknowledge J. Adrover, V. Lozano, P. Stark and T. Speed for helpful discussions and/or encouragement. This work was partially supported by NSF Grant DMS-9404276.

7. Appendix: Proofs

Pf.1. Start with the sufficient statistic $\mathbf{P} = (P_2, ..., P_L)$, $L > 3$. A recursive relation between N_l and N_{l+1} leads to two familes of unbiased estimators of zero

$$V_l(\mathbf{P}) = (2l+5)P_lP_{l+1} - (2l+3)(2l+5)P_lP_{l+2} + (2l+3)P_{l+1}^2 - (2l+5)P_{l+1}P_{l+2},$$

$$U_l(\mathbf{P}) = 4P_lP_{l+2} - 4(l+2)P_lP_{l+3} - 4P_{l+1}P_{l+3} + 4(l+2)P_{l+1}P_{l+2}.$$

Unbiased estimators of zero satisfying the scaling conditions (4) can be obtained easily from these. We use the V_l to show that $\hat{Q}_0^2$ is not UMVU if $L > 3$. For an unbiased estimator to be UMVU it is necessary and sufficient that it be uncorrelated with any unbiased estimator of zero. Since the identity

$$5 < V_2, P_2 >= 18Q^6\,[\,N_3(n) - 7N_4(n)\,] = 0$$

can not hold for all n, it follows that $\hat{Q}_0^2$ is not UMVU when $L > 3$. A similar reasoning shows that no P_l is an UMVU estimator of its expectation.

Pf.2. Take a_{lm}, $l \leq L$, so $\mathbf{P} = (P_2, ..., P_L)$. We show that an UMVU estimator of n can not be independent of P_l, P_{l+1}, for some $l \leq L-1$. Suppose $\hat{n}(\mathbf{P})$ is an unbiased estimator of n that is independent of P_l and P_{l+1}. Then

$$E(P_{l+1}) = Q^2 N_{l+1}(n) = Q^2 \frac{2l-1+n}{2l+5-n} N_l(n).$$

Since $\hat{n}$ is an unbiased estimator for any Q and n, we have

$$E(P_{l+1})E(2l+5-\hat{n}) = Q^2(2l-1+n)N_l(n) = E(2l-1+\hat{n})E(P_l).$$

But P_l and P_{l+1} are independent of $\hat{n}$ by assumption, so

$$E[\,(2l+5-\hat{n})P_{l+1} - (2l-1+\hat{n})P_l\,] = 0,$$

for all Q and n. That is,

$$U = (2l+5-\hat{n})P_{l+1} - (2l-1+\hat{n})P_l$$

is an unbiased estimator of zero. If in addition $\hat{n}$ is minimum variance, then it must be uncorrelated with any unbiased estimator of zero. In particular we must have $< U, \hat{n} >= 0$, thus

$$N_l(n) + N_{l+1}(n) = 0.$$

This is a contradiction since the last identity is not true for all n. Therefore $\hat{n}$ is not an UMVU estimator of n.

Pf.3. Consider the parameter space $-3 < n < 3$, $Q > 0$, and assume P_2 is excluded from the data, i.e. $\mathbf{P} = (P_3, ..., P_L)$. We show that no unbiased estimator of Q^2 satisfying the scaling conditions (4) exists. Take a finite variance unbiased estimator $\delta(\mathbf{P})$ of Q^2 , so that

$$E(\delta) = Q^2 \tag{7}$$

for any n and Q. If $\delta(\mathbf{P})$ satisfies the scaling conditions, then

$$\delta(\mathbf{P}) = P_3 F(\frac{P_4}{P_3}, ..., \frac{P_L}{P_3}),$$

for some function F. Also, by the scaling conditions it suffices to consider the case $Q = 1$. Define new variables $Y_1 = P_3, Y_2 = P_4/P_3, ..., Y_{L-2} = P_L/P_3$. The joint density of the Y_l is

$$f(Y_1, ..., Y_{L-2}) \propto \frac{Y_1^{\frac{L^2+2L-10}{2}} Y_2^{\frac{7}{2}} \cdots Y_{L-2}^{\frac{2L-1}{2}} \exp(-Y_1[\frac{7}{2N_3(n)} + \cdots + \frac{(2L+1)Y_{L-2}}{2N_L(n)}])}{N_3(n)^{\frac{7}{2}} N_4(n)^{\frac{9}{2}} \cdots N_L(n)^{\frac{2L+1}{2}}}.$$

Calculating the integral with respect to Y_1 in the expectation of δ turns Eqn.(7) into

$$\int \frac{\Theta(\mathbf{Y}')d\mathbf{Y}'}{P(n, \mathbf{Y}')^{\alpha}} = kN_3(n)^{-1}\bar{N}_4(n)^{\frac{9}{2}} \cdots \bar{N}_{L-2}(n)^{\frac{2L+1}{2}}, \tag{8}$$

where $\mathbf{Y}' = (Y_2, ..., Y_{L-2})$, $\alpha = (L^2 + 2L - 6)/2$, k is some positive constant and

$$\bar{N}_l(n) = \frac{N_l(n)}{N_3(n)} = \frac{n+5}{11-n} \cdots \frac{n+2l-3}{2l+3-n},$$

$$P(n, \mathbf{Y}') = 7 + \frac{9}{\bar{N}_4(n)} Y_2 + \cdots + \frac{2L+1}{\bar{N}_L(n)} Y_{L-2},$$

$$\Theta(\mathbf{Y}') = Y_2^{\frac{7}{2}} \cdots Y_{L-2}^{\frac{2L-1}{2}} F(\mathbf{Y}').$$

We first proof that the integral on the left side of Eqn.(8) is bounded for any n: since the variance of δ is finite for any n, it follows that $E(|\ \delta\ |) < \infty$ for all n. Choose $n_0 = 1$, then for any n

$$\left| \int \frac{\Theta(\mathbf{Y}') d\mathbf{Y}'}{P(n, \mathbf{Y}')^\alpha} \right| \leq \int \frac{|\Theta(\mathbf{Y}')|}{P(n_0, \mathbf{Y}')^\alpha} \frac{P(n_0, \mathbf{Y}')^\alpha}{P(n, \mathbf{Y}')^\alpha} d\mathbf{Y}'. \tag{9}$$

The integral of the first quotient on the right is bounded by assumption. The second quotient is a ratio of two positive polynomials of the same degree. It is easy to see that this ratio is bounded and therefore the integral on the left side of Eqn.(9) is bounded. But, since $N_3(n) \propto (3+n)/(9-n)$, the right hand side of Eqn.(8) goes to infinity as $n \to -3$. A contradiction since the left side is bounded.

Pf.4. It was shown in Section 5 that the correlation sign of efficient estimators of n and Q is determined by the sign of I_{12}. We use the same argument for different spectrum normalizations. Renormalize the power spectrum to the kth multipole so that

$$E(P_l) = R^2 M_l(n),$$

where $M_k(n) = 1/(2k+1)$ and $(2k+1)M_l(n) = N_l(n)/N_k(n)$. The new information matrix $I'(R^2, n)$ is the same as the previous $I(Q^2, n)$ in Section 5, but with R in place of Q and M_l instead of N_l. Writing I'_{12} in terms of N_l leads to

$$I'_{12} = \frac{1}{2R^2} \sum_{l \leq L} (2l+1)\, \partial_n ln(M_l) = \frac{1}{2R^2} \sum_{l \leq L} (2l+1)[\, \partial_n ln(N_l) - \partial_n ln(N_k)\,],$$

which can be re-written as

$$I'_{12} = -\frac{N}{2R^2} [\, \partial_n ln(N_k) \;-\; \frac{1}{N} \sum_{l \leq L} (2l+1) \partial_n ln(N_l)\,],$$

where $N = \sum_{l \leq L}(2l+1)$. That is, the term I'_{12} is positive when $\partial_n ln(N_k)$ is to the left of the weighted average $\frac{1}{N}\sum_{l \leq L}(2l+1)\partial_n ln(N_l)$ and negative otherwise.

References

Backus, G. (1970) Inference form Inadequate and Inaccuarate Data I-III. *Proc. Nat. Acad. Sci.*, **65**, pp.1-7; **65**, pp.281-287; **67**, pp.282-289

Backus, G. (1988) Comparing Hard and Soft Prior Bounds in Geophysical Inverse Problems. *Geophys. J.*, **94**, 249-261

Backus, G. (1989) Confidence Set Inference with a Prior Quadratic Bound. *Geophys. J.*, **97**, 119-150

Banday,A., Gorsky, K., Tenorio, L., Wright, E.L., Smoot, G., Lineweaver, C., Kogut, A., Hinshaw, G. & Bennett, C. (1994) On the RMS Anisotropy at 7° and 10° Observed in the COBE-DMR Two Year Sky Maps. *Ap. J.*, **436**, L99-L102

Bennett *et al.* (1994) Cosmic Temperature Fluctuations from Two Years of COBE Differential Microwave Radiometer Observations. *Ap. J.*, **436**, 423-442

Benton, E.R., Estes, R.H., Langel, R.A. & Muth, L.A. (1982) Sensitivity of Selected Geomagnetic Properties to Truncation Level of Spherical Harmonic Expansions. *Geophys. Res. Lett.*, bf 9, 254-257

Bond, J.R. & Efstathiou, G. (1987) The Statistics of Cosmic Background Radiation Fluctuations. *MNRAS*, **226**, 655-687

Cayón, L. & Smoot, G. (1995) Host and cold Spots in the First Plus Secod Year COBE-DMR Maps. *Ap. J.* In Press

Efstathiou, G. (1990) Cosmological Perturbations. In *Physics of the Early Universe: Proc. 36th Scottish Universities Summer School in Physics* (eds. Peacock, J. Heavens, A.F. & Davies, A.T.) pp.361-463

Górski, K., Hinshaw, G., Banday, A., Bennett, C., Wright, E., Kogut, A., Smoot, G. & Lubin, P. (1994) On Determining the Spectrum of Primordial Inhomogeneity from the COBE DMR Sky Maps: Results of Two Year Data Analysis. *Ap. J.*, **430**, L89-L92

Hoadley, B. (1971) Asymptotic Properties of Maximum Likelihood Estimators for the Independent not Identically Distributed Case. *Ann. Math. Statist.*, **42**, 1977-1991

Kogut, A., Banday, A., Bennett, C., Hinshaw, G., Lubin, P.M. & Smoot, G. (1995) Gaussian Statistics of the COBE Microwave Background: Correlation of Temperature Extrema in the COBE DMR Two-Year Sky Maps. *Ap. J.*, **439**, L29-L32

Lineweaver, C. H. (1994) *Correlation Function Analysis of the COBE Differential Microwave Radiometer Sky Maps*, Ph.D Thesis, Berkeley

Lineweaver, C. H. (1995) Recent COBE Results. In *Proc. 30th Rencontres de Moriond on Dark Matter in Cosmology, Clocks and Tests of Fundamental Laws* (ed. J. Tran Thanh Van), Editions Frontieres. In press

Penzias, A. A., Wilson, R. W. 1965, *Ap. J.*, **142**, 419

Seljak, U. & Bertschinger E. (1993) Maximum-Likelihood Analysis of the COBE Angular Correlation Function. *Ap. J.*, **417**, L9-L12

Smoot, G.F. *et al.* (1991) Preliminary Results from the COBE Differential Microwave Radiometers: Large Angular Scale Isotropy of the Cosmic Microwave Background. *Ap. J.*, **371**, L1-L5

Smoot, G.F. *et al.* (1992) Structure in the COBE DMR First Year Maps. *Ap. J.*, **396**, L1-L5

Smoot, G.F., Tenorio, L., Banday, A., Kogut, A., Wright, E., Hinshaw, G., Bennett, C. (1994) Statistics and Topology of the COBE DMR First Year Sky Maps. *Ap. J.*, **437**, 1-11

Stark, P.B. (1992) Minimax Confidence Intervals in Geomagnetism. *Geophys. J. Int.*, **108**, 329-338

Stark, P.B. (1993) Uncertainty of the Cosmic Background Explorer Quadrupole Detection. *Ap. J.*, **408**, L73-L76

Wright, E.L., Smoot, G.F., Kogut, A., Hinshaw, G., Tenorio, L., Lineweaver, C., Bennett, C. and Lubin, P.M. (1994) Comments on the Statistical Analysis of Excess Variance in the COBE Differential Microwave Radiometer Maps. *Ap. J.*, **420**, 1-8

This manuscript was prepared with the AAS LaTeX macros v3.0.

Figure Captions

Figure 1. Estimates of Power Spectrum Parameters n and Q (also known as Q_{rms-PS}). For many of the points Q was found after fixing $n = 1$. The curves are estimates of the conditional expectation of $\hat{Q}$ given $\hat{n}$. Only the five points on the lower right were calculated with the quadrupole included in the data (Lineweaver 1995).

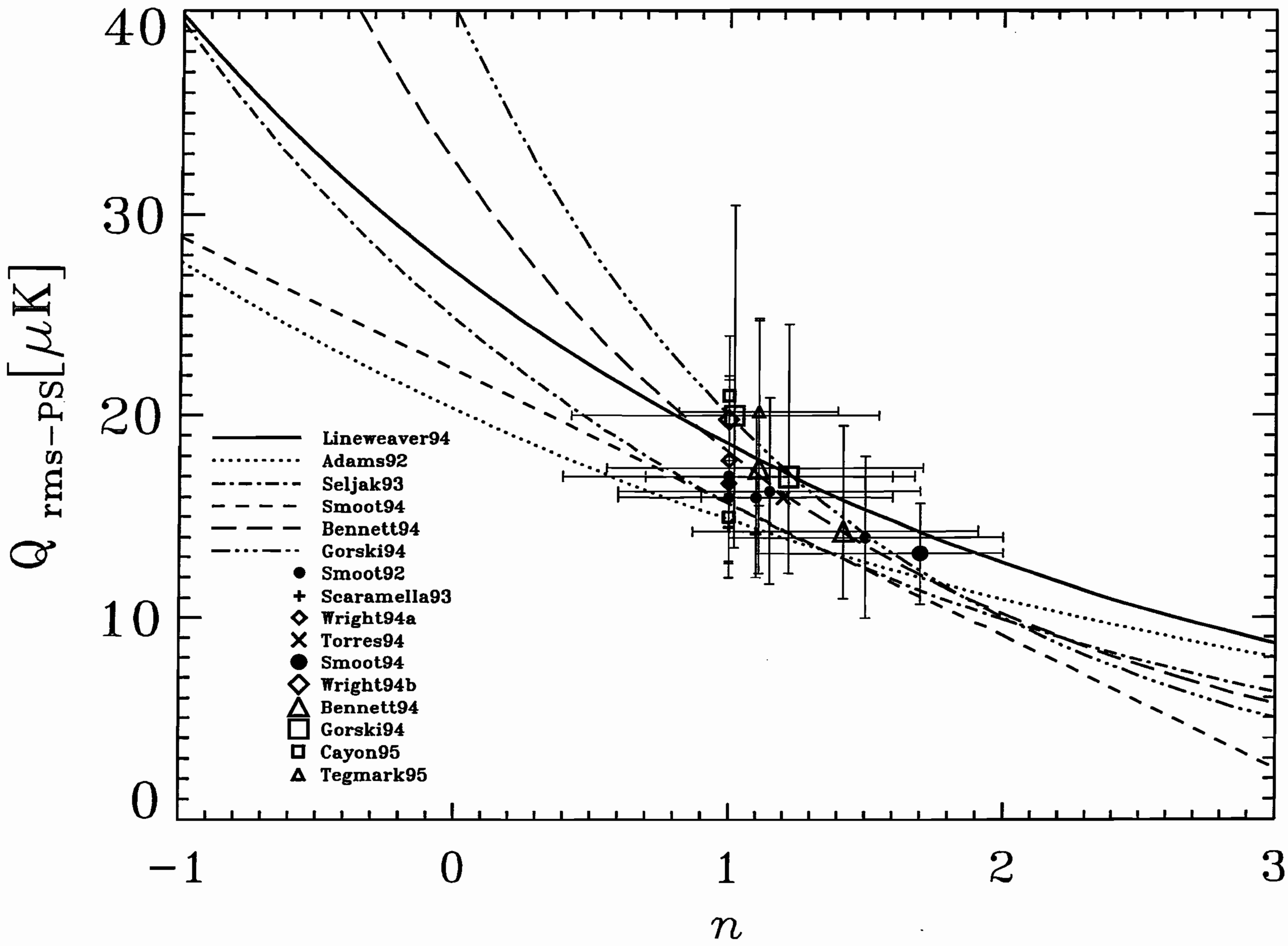
Figure 2. Scatter plot of simulated ML estimators of Q^2 and n when only P_2 and P_3 are used. The input values were $n_0 = 1$ and $Q_0 = 1$. The dotted curve is the conditional expectation of $\hat{Q}^2_{\rm ml}$ given $\hat{n}_{ml}$.

Figure 3. Bias of the ML estimators for different values of n with $Q = 1$. Solid symbols correspond to simulations with the quadrupole excluded.

Figure 4. Comparing the standard error of $\hat{n}_{\rm ml}$ (4a) and $\hat{Q}^2_{\rm ml}$ (4b) with the square root of their CR bound (curves) as functions of n with $Q_0 = 1$. The solid symbols correspond to the cases when the quadrupole is excluded.

Figure 5. Least-squares fits to a 2-year 53 GHz COBE sky map. Estimates of the x-component of the dipole in celestial coordinates and standard errors are shown in 5a and 5b respectively. The fits were done to orders $L = 1, 2, 4, ..., 12$, and regions with galactic latitudes $|b| < 5°$ and $10°$ were removed to minimize contamination from the galactic plane.

Figure 1



40
30
20
10
0
−1
0
1
2
3
n
Q rms−PS[μK]
Lineweaver94
Adams92
Seljak93
Smoot94
Bennett94
Gorski94
Smoot92
Scaramella93
Wright94a
Torres94
Smoot94
Wright94b
Bennett94
Gorski94
Cayon95
Tegmark95

Figure 2

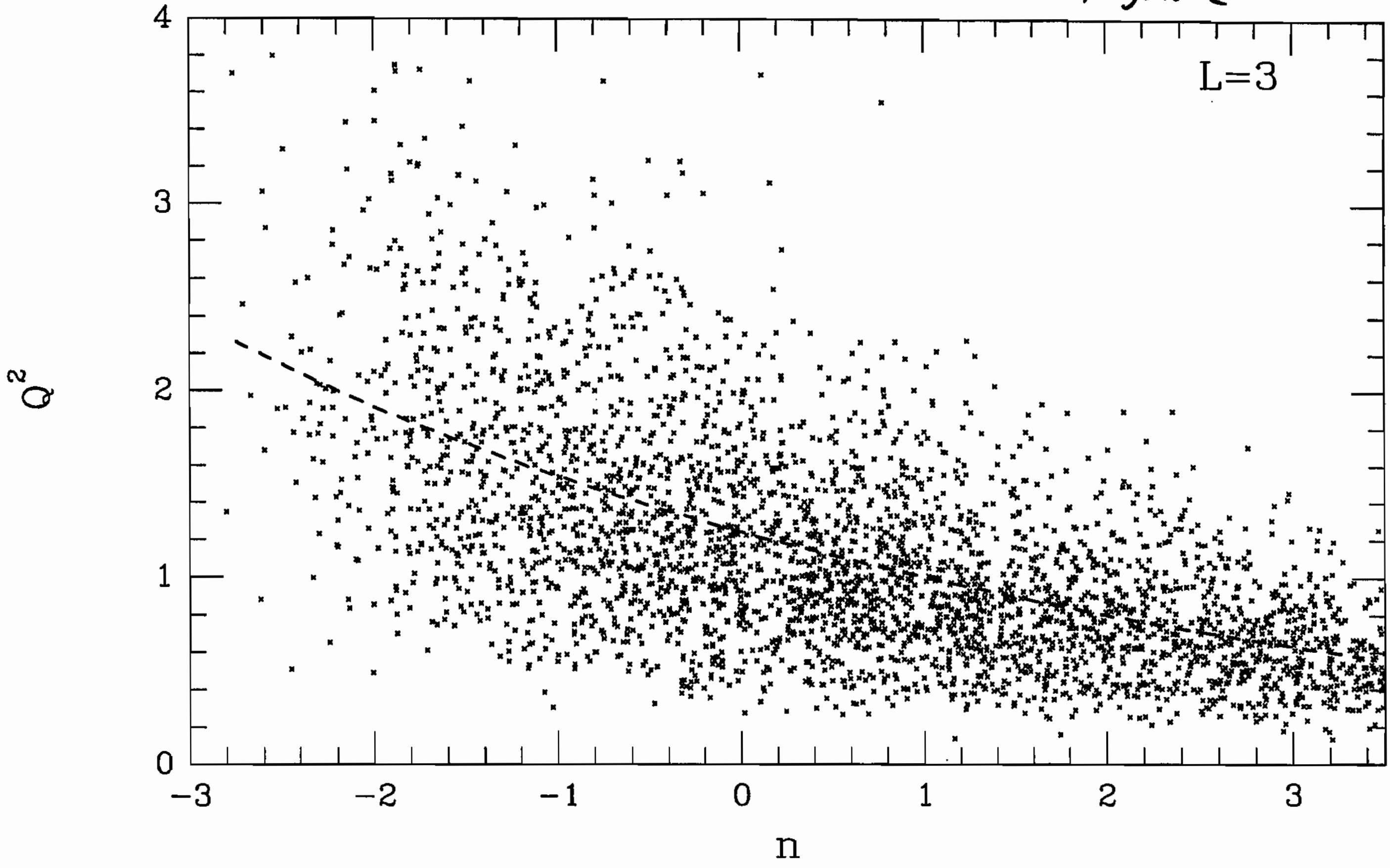


Figure 3a

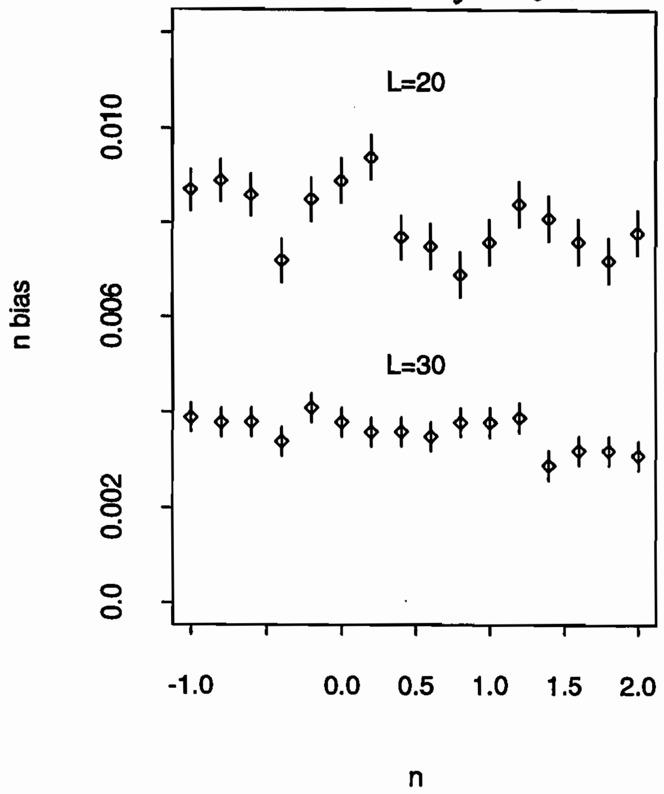


Figure 3b

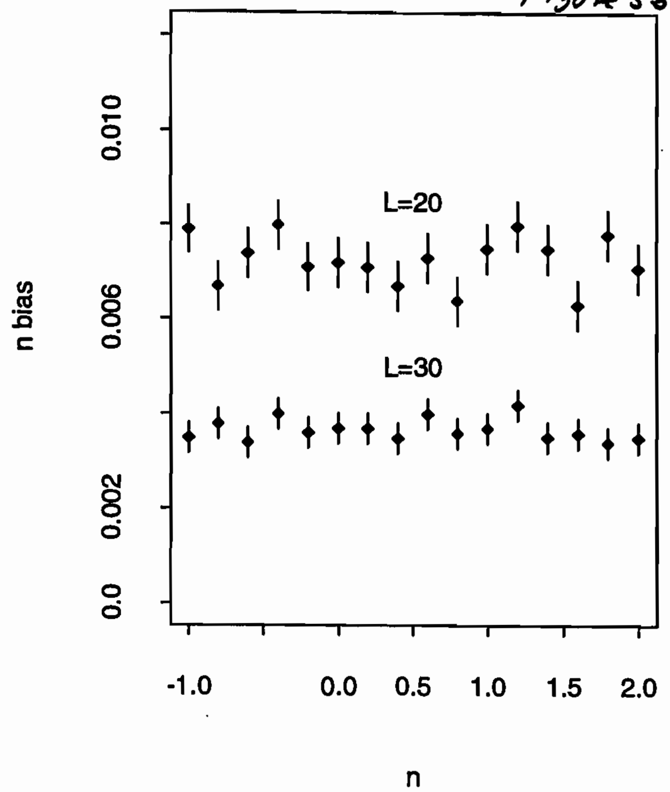


Figure 3c

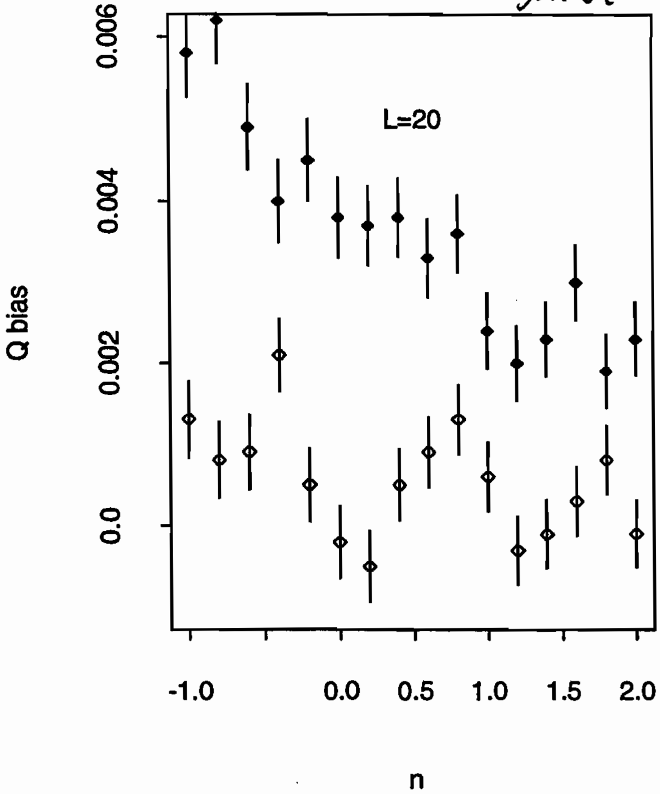


Figure 3d

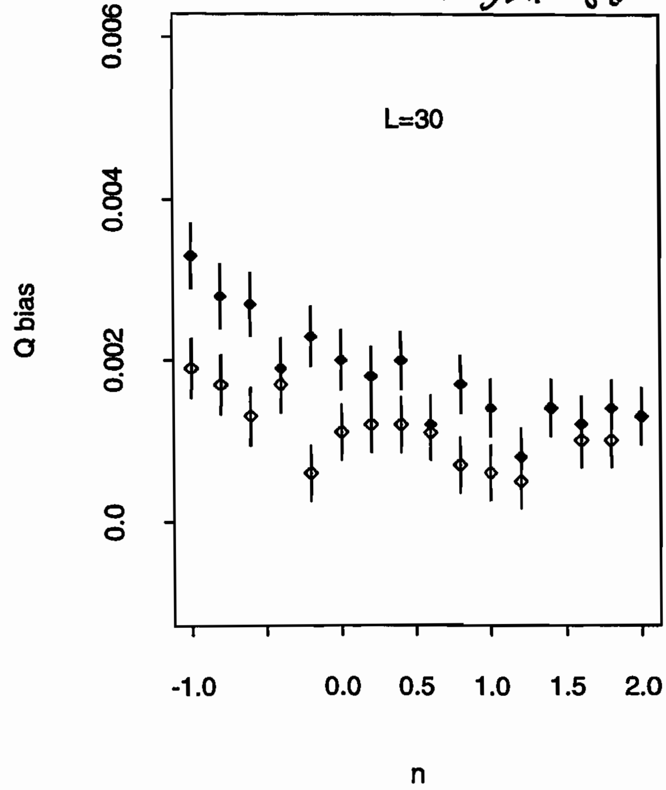


Figure 4a

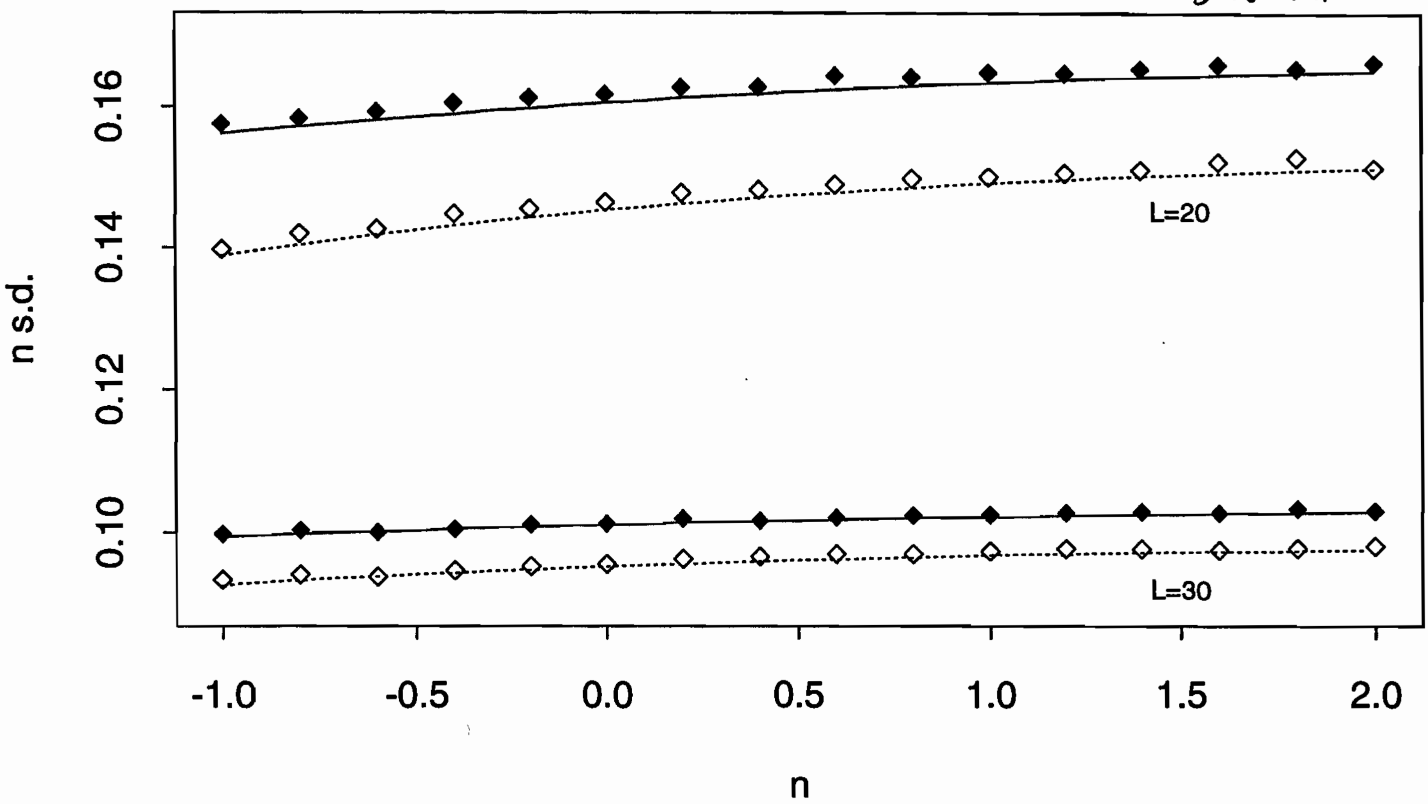


Figure 4b

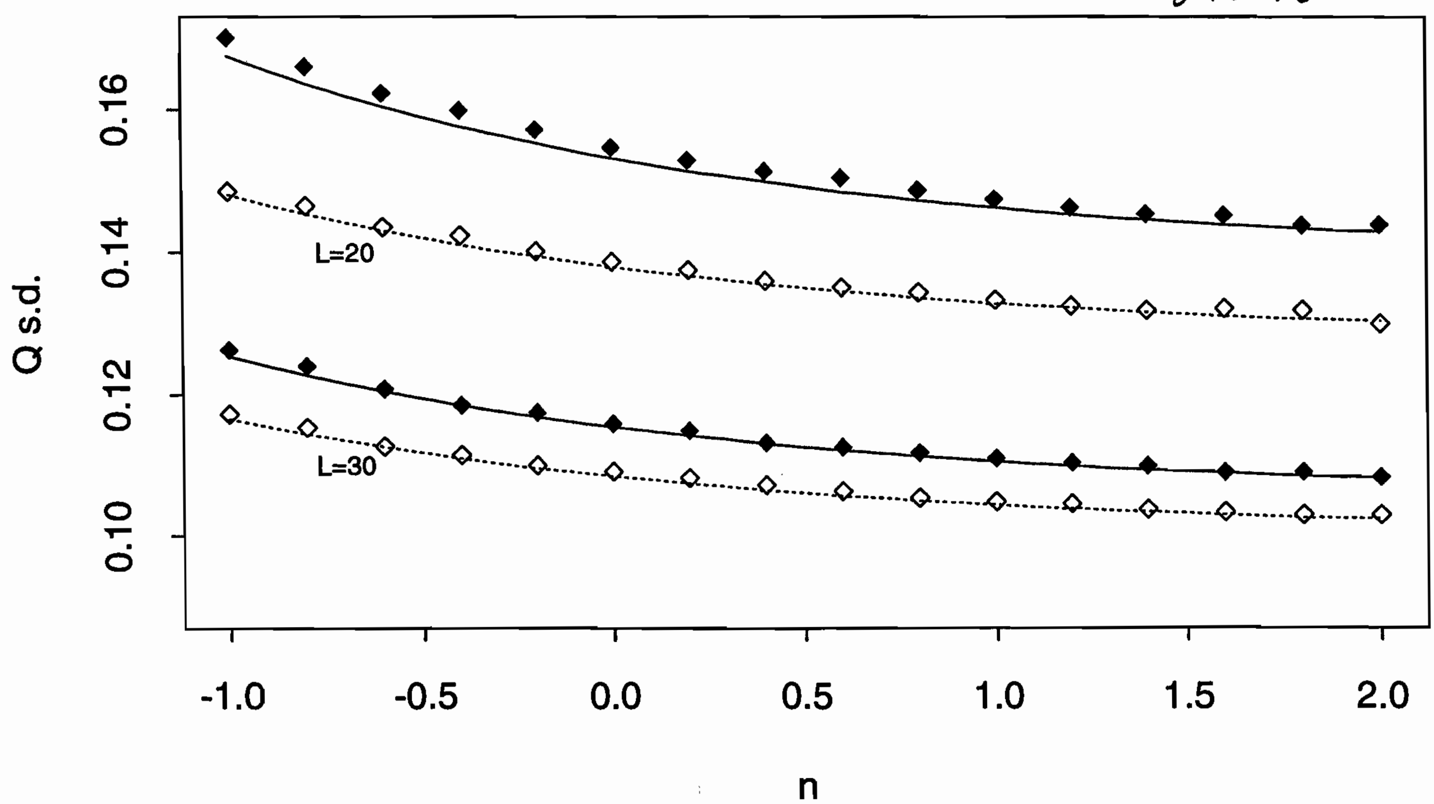


Figure 5a

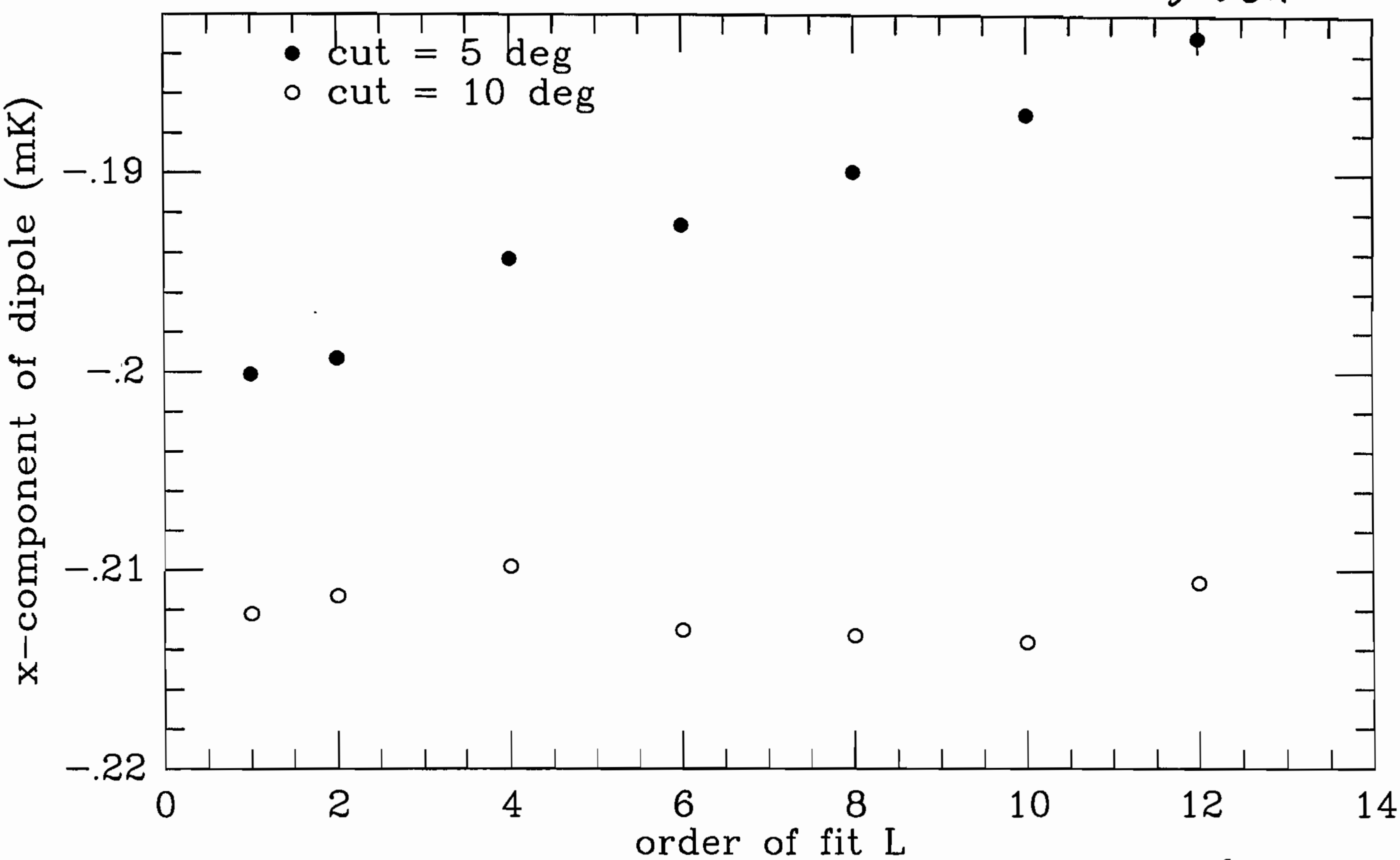


Figure 5b