Cosmic Microwave Background Anisotropy Window Functions Revisited

Lloyd Knox

(August 17, 2001)

The primary results of most observations of cosmic microwave background (CMB) anisotropy are estimates of the angular power spectrum averaged through some broad band, called band-powers. These estimates are in turn what are used to produce constraints on cosmological parameters due to all CMB observations. Essential to this estimation of cosmological parameters is the calculation of the expected band-power for a given experiment, given a theoretical power spectrum. Here we derive the "band-power" window function which should be used for this calculation, and point out that it is not equivalent to the window function used to calculate the variance. This important distinction has been absent from much of the literature: the variance window function is often used as the band-power window function. We discuss the validity of this assumed equivalence, the role of window functions for experiments that constrain the power in *multiple* bands, and summarize a prescription for reporting experimental results. The analysis methods detailed here are applied in a companion paper to three years of data from the Medium Scale Anisotropy Measurement.

I. INTRODUCTION

Measurement of the anisotropy of the Cosmic Microwave Background (CMB) is proving to be a powerful cosmological probe. However, exact statistical treatment of the data is complicated and time-consuming, and promises to become prohibitively so in the very near future. This difficulty explains why no one has calculated the likelihood of cosmological parameters, given the available data from all CMB experiments.

Instead, constraints on cosmological parameters have been derived by approximate methods—namely, the use of "radical compression" [1]. Reduction of CMB data to estimates of the angular power spectrum, C_l, can be viewed as a form of data compression. This compressed data is then in turn used to constrain cosmological parameters. Usually, the compression is not to estimates of the individual C_ls themselves, but to band-powers [2] which are averages of the power spectrum through a certain filter, or window function.

These band-powers, together with their window functions, have traditionally been the main results of CMB experiments. Unfortunately, a large number of experimental results papers only provide what we will call the variance window function, and not the "band-power" window function. Indeed, the distinction between the two has been missing in much of the literature. They are not equivalent, except in the limit of vanishing off-diagonal signal correlations. Reports of constraints on the CMB power spectrum should contain the latter type of window function, together with the quantification of the uncertainty in the band-power.

Not all reductions of CMB data to power spectrum estimates have been presented with only variance window functions. As a rough guide, those that have been analyzed with "quadratic estimators" have the right (band-power) window function while those that have been analyzed by evaluation of the likelihood function do not. The persistence of this confusion is probably due to the fact that likelihood analysis obscures the relation between the data and the derived band-power, which is much clearer when one uses an estimator. It has been shown [3] that a particular quadratic estimator [3–5] is guaranteed to produce the maximum-likelihood result (if used iteratively) and below we exploit this fact to derive the expression for the band-power window function appropriate for likelihood analysis.

In a companion paper [6] we apply the radical compression procedures detailed here to three years of data from the Medium Scale Anisotropy Measurement (MSAM). This procedure is a combination of techniques developed in [3], [1] and here. The application to MSAM strongly demonstrates the power of this method compared to the usual approach of compression to flat band-powers. In particular, the analysis in the companion paper results in a new and significant constraint at $l \sim 400$, a theoretically very interesting region of the power spectrum, which had been previously obscured by use of variance window functions, rather than band-power window functions.

II. THE BAND-POWER METHOD

A very useful meeting point for theory and experiment is provided by the band-power. An important property of a meeting point, is that both parties planning on meeting should be able to get there. Although the directions for going from the data to the band-power are already clearly explained in the literature (and will be reviewed below), those for going from the theory to the band-power are not. The point of this paper is to provide those directions — directions which are clearly essential to the confrontation.

We now explicitly define the band-power method [2] which has been used by many authors, *e.g.*, [1,7,8]. In the simplest case of a dataset that produces one band-power, its calculation is conceptually straightforward: the power spectrum is assumed to be flat ($\mathcal{C}_l$ independent of l, where $\mathcal{C}_l \equiv \frac{l(l+1)}{2\pi} C_l$) and the band-power estimate is taken to be the amplitude of $\mathcal{C}_l$ estimated from the data (*e.g.*, via likelihood analysis). For dataset B, let us call this maximum likelihood value, $\mathcal{C}_B$. Let us further assume that the uncertainty in $\mathcal{C}_B$ is Gaussian-distributed with variance σ_B^2.

Because theoretical power spectra are not flat, the relation between a theoretical power spectrum and the prediction of that theory for the $\mathcal{C}_B$ derived from data is non-trivial. The theoretical prediction is simply the expectation value of $\mathcal{C}_B$, given that the theory is true. Since $\mathcal{C}_B$ is a determination of the amplitude of the power spectrum we will assume a linear dependence of its expectation value on the power spectrum of the theory, specified by the band-power window function, W_l^B:

$$\langle \mathcal{C}_B \rangle = \sum_l \left(W_l^B / l \right) \mathcal{C}_l(a_p) \tag{1}$$

where a_p are the parameters of the theory. These parameters, a_p, could be cosmological parameters (*e.g.*, Ω_b, Ω_Λ, H_0, etc.) or parameters from a phenomenological power spectrum. Throughout we will assume that W_l^B is normalized so that $\langle \mathcal{C}_B \rangle = \mathcal{C}_l$ for $\mathcal{C}_l$ independent of l.

With the assumptions of independence and Gaussianity for the uncertainty in $\mathcal{C}_B$ and the specification of the linear relationship between $\mathcal{C}_l$ and $\langle \mathcal{C}_B \rangle$, it follows that the likelihood of the parameters is maximized by minimizing the following χ^2:

$$\chi^2 = \sum_B \Big(\sum_l \left(W_l^B / l \right) \mathcal{C}_l(a_p) - \mathcal{C}_B \Big)^2 / \sigma_B^2, \tag{2}$$

This χ^2 represents the confrontation between theory and data that occurs at the meeting point of the band-power. Use of this equation, or ones similar to it, is very efficient and is what has been used in analyses of the constraints placed on parameters due to available CMB data. Note that W_l^B projects the theory into the "plane" of the experiment — or rather the same plane into which the experimental data have also been reduced.

Previous work has focused on generalizing Eq. 2 to take into account the non-Gaussianity and dependence of the uncertainties in $\mathcal{C}_B$ [1,7]. Here we focus on the choice of W_l^B. It is often assumed to be equal to the variance window function, W_l^V, which is actually the diagonal element of a window function *matrix* which specifies the relationship between the angular power spectrum and the covariance matrix of the signal, S:

$$S_{pp'} \equiv \langle s_p s_{p'} \rangle = \sum_l \mathcal{C}_l W_{l,pp'} / l. \tag{3}$$

where s_p is the signal contribution to the pth element of a dataset and the brackets indicate ensemble average. For a single demodulation, all the diagonal elements of the window function matrix are equal and that is why we can speak of *the* diagonal element.

While using the variance window function to calculate the signal covariance matrix is correct (this is what the variance window function is defined to do, see, *e.g.*, [9–11]), using it in Eq. 2 is not (except in the special case specified below).

It is perhaps worth emphasizing the prevalent use of the variance window functions in equations like Eq. 2. All of the references in ref. [1,6] use it, as do all published analyses of large numbers of band powers. This use of W_l^V is due to the fact that a large number of reports of band-power constraints do not include W_l^B, but only W_l^V (e.g. [12]).

III. THE BAND-POWER WINDOW FUNCTION

A minimum-variance, unbiased, estimate of the power spectrum is given by [3,4]

$$\mathcal{C}_l = \frac{1}{2} F_{ll'}^{-1} \mathrm{Tr} \left[(\Delta\Delta - N) (S+N)^{-1} \frac{\partial S}{\partial \mathcal{C}_{l'}} (S+N)^{-1} \right] \tag{4}$$

where

$$F_{ll'} = \mathrm{Tr} \left[(S+N)^{-1} \frac{\partial S}{\partial \mathcal{C}_l} (S+N)^{-1} \frac{\partial S}{\partial \mathcal{C}_{l'}} \right] \tag{5}$$

is called the Fisher matrix.

If we are only interested in estimating the amplitude of a power spectrum that we assume to be flat, ($\mathcal{C}_l = \mathcal{C}_B = \mathrm{constant}$) we can rewrite the minimum-variance estimator for $\mathcal{C}_l$ (Eq. 4) as

$$\mathcal{C}_B = \frac{1}{2} F_{BB}^{-1} \mathrm{Tr} \left[(\Delta\Delta - N) (S+N)^{-1} \sum_{l'} \frac{\partial S}{\partial \mathcal{C}_{l'}} (S+N)^{-1} \right] \tag{6}$$

where $F_{BB} = \sum_{ll'} F_{ll'}$ (because $\partial S / \partial \mathcal{C}_B = \sum_l \partial S / \partial \mathcal{C}_l$). Equation 6 can be viewed as the directions that take one from the data, Δ, to the meeting point of the band-power.

We now must provide the directions to go from a theory to the band-power. With the usual assumptions of Gaussianity and statistical isotropy, theories are completely specified by their angular power spectrum, $\mathcal{C}_l$. We are therefore after the expectation value of the $\mathcal{C}_B$ of Eq. 6, under the assumption that the true power spectrum is $\mathcal{C}_l$. Calculation of the dependence of this expectation value on $\mathcal{C}_l$ will provide the directions we need.

The expectation value is easily calculated after noting that $\langle (\Delta\Delta - N) \rangle = S = \sum_l \mathcal{C}_l \left(\frac{\partial S}{\partial \mathcal{C}_l} \right)$; it is given by

$$\langle \mathcal{C}_B \rangle = \frac{1}{2} F_{BB}^{-1} \mathrm{Tr} \left[\sum_{ll'} \mathcal{C}_l \frac{\partial S}{\partial \mathcal{C}_l} (S+N)^{-1} \frac{\partial S}{\partial \mathcal{C}_{l'}} (S+N)^{-1} \right] \tag{7}$$

which further simplifies to

$$\langle \mathcal{C}_B \rangle = \frac{\sum_{ll'} \mathcal{C}_l F_{ll'}}{\sum_{ll'} F_{ll'}} = \sum_l \left(W_l^B / l \right) \mathcal{C}_l \tag{8}$$

which implicitly defines the band-power window function:

$$W_l^B / l = \frac{\sum_{l'} F_{ll'}}{\sum_{ll'} F_{ll'}}. \tag{9}$$

We have found our linear relationship between the expected value of $\mathcal{C}_B$ and the assumed power spectrum, $\mathcal{C}_l$. We note that it has a form we might have guessed — an inverse-variance weighted sum of the $\mathcal{C}_l$. We can identify it as such because the Fisher matrix also serves as an approximation to the inverse of the covariance matrix of the uncertainty in the $\mathcal{C}_l$ estimates.

For band-powers derived from an estimator, the derivation of the band-power window function is quite straightforward: one simply calculates the expectation value, given $\mathcal{C}_l$, as done above. However, band-powers are often determined instead by finding the maximum of the likelihood function, rather than by the quadratic estimator of Eq. 4. The maximum-likelihood estimate is a complicated, non-quadratic, function of the data and its expectation value is not easy to calculate. In fact, for a maximum-likelihood estimate the band-power window function is ill-defined because the relationship between the maximum-likelihood and $\mathcal{C}_l$ is non-linear.

Nevertheless, the above expressions for the band-power window function are still useful for maximum-likelihood band-power estimates. This is due to the relationship between likelihood analysis and the quadratic estimator of Eq. 4 pointed out in [3]: used iteratively, Eq. 4 results in the band-power that maximizes the likeliihood. If the $\mathcal{C}_l$ assumed for the right-hand side of Eq. 4 is at least roughly consistent with the data (and "smooth" [1,3,13,14]) then a single iteration of Eq. 4 will produce a very good approximation to the maximum-likelihood. Thus, the band-power window function is appropriate for $\mathcal{C}_l$ sufficiently close to the most likely power spectrum. Further away, non-linear corrections will become important. One could, in principle, calculate these non-linear corrections, but the improved precision is probably not worth the additional complication.

We emphasize that Eq. 8 truly does specify a linear relationship between the quadratically estimated, $\langle \mathcal{C}_B \rangle$, and $\mathcal{C}_l$. One might suspect that there are other dependencies on $\mathcal{C}_l$ hidden in the window function itself. However, the Fisher matrix that appears twice in Eq. 9 is that for a flat power spectrum with amplitude $\mathcal{C}_B$, and does not depend on $\mathcal{C}_l$.

IV. THREE EXAMPLES

Consider an experiment that maps the whole sky with a Gaussian beam with full width at half max $= \sqrt{8 \ln 2}\sigma_b$ and a uniform noise level specified by a weight-per-solid angle, w. In this case of uniform noise and full-sky coverage, the Fisher matrix can be calculated analytically and is given by [15]

$$F_{ll'} = \frac{2l+1}{2} \left[\mathcal{C}_l + \frac{l(l+1)}{2\pi w B^2(l)} \right]^{-2} \delta_{ll'} \tag{10}$$

where $B(l) = e^{-l^2 \sigma_b^2 / 2}$. Due to the $\delta_{ll'}$, the sum over l' is trivial and the band-power window function is

$$W_l^B / l \propto \frac{2l+1}{2} \left[\mathcal{C}_l + \frac{l(l+1)}{2\pi w B^2(l)} \right]^{-2}. \tag{11}$$

Thus, for $\mathcal{C}_l$ constant, we see that the band-power window function for this map is proportional to l^2 at low l and then eventually drops very rapidly at higher l where it is proportional to $B^4(l)/l^2$.

This behavior of W_l^B is intuitively reasonable. Cosmic variance is the reason that the very low ls are less important to the overall determination of the band-power, and instrument noise suppresses the importance of the very high ls.

Contrast this behavior to that of the variance window function. For a map, W_l^V is simply given by the square of the spherical harmonic transform of the beam, $B(l)$. Therefore

$$W_l^V \propto B^2(l). \tag{12}$$

Note that this implies that the most important moments are the ones at lowest l! Further, there is no dependence on the noise level. For the band-power window function we see that as the noise is lowered (w raised), the importance of the higher l moments increases.

Our second example is for a dataset with n points that have no signal or noise correlations. We leave it as an exercise for the reader to show that in this case, $W_l^B = W_l^V$. That is, in the absence of correlations, the two window functions are equivalent.

Our final example is from the Medium Scale Anisotropy Measurement (MSAM) 3-year dataset [6]. This dataset was reduced to measurements of the sky with two different beam maps, called single-difference and double-difference. The high signal-to-noise and dense sampling of the dataset mean that it is sensitive to $\mathcal{C}_l$ at somewhat higher values of l then one would infer from the variance window function. See Fig. 1. For the single-difference measurements, the results are especially striking: the peak is shifted from $l = 120$ to $l = 160$ and at $l = 400$, where there is a second local peak, W_l^B is about 5 times larger than W_l^V. The band-power window functions were calculated assuming a flat power spectrum with amplitude consistent with the data of $\mathcal{C}_B = 2000(\mu\mathrm{K})^2$.

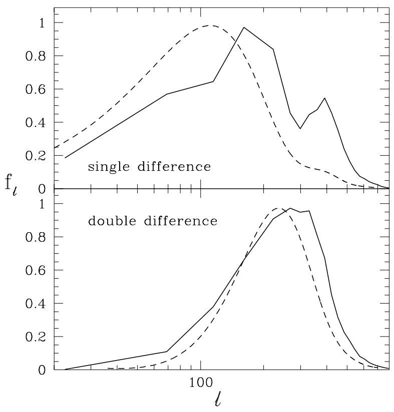

FIG. 1. The band-power window functions (solid lines) and variance window functions (dashed lines) for the single and double difference MSAM beam maps ($f_l \equiv W_l/l$). The normalization here is arbitrary.

V. MULTIPLE BANDS

In the above we have assumed a flat power spectrum. This is overly restrictive given that we wish to determine the presence of features in this power spectrum! However, the above easily generalizes to the case where the power spectrum is determined in multiple flat bands, each of finite extent in l.

Some datasets have sufficient dynamic range to estimate the power spectrum in more than one band. For these datasets we can parameterize the power spectrum, $\mathcal{C}_l$, with the power in bands enumerated by the subscript B:

$$\mathcal{C}_l = \sum_B \chi_{B(l)} \mathcal{C}_B. \tag{13}$$

Where $\mathcal{C}_B$ is the amplitude of the power spectrum within band B and

$$\chi_{B(l)} = \left\{ \begin{array}{lcl} 1 & : & l_<(B) < l < l_>(B) \\ 0 & : & \text{otherwise} \end{array} \right\} \tag{14}$$

where $l_<(B)$ and $l_>(B)$ delimit the range of band B. For each of these bands we can calculate a band-power window function via

$$W_l^B/l = \frac{\sum_{l'} F_{ll'}}{\sum_{ll'} F_{ll'}} \tag{15}$$

where the sums over l' run only from $l_<(B)$ to $l_>(B)$.

We could remove the need for window functions by making the bands very narrow since sufficiently narrow bands ensure that the sensitivity to each $\mathcal{C}_l$ within the band is approximately independent of l. However, making the bands too narrow makes the error bars very large and highly correlated. This is undesirable for two reasons. First, it hinders visual interpretation of the results and second, the larger the error bars, the more important are the non-Gaussian aspects of the distribution. And although the ansatzes of [1] for this non-Gaussian distribution have been shown to work quite well in some cases, it is not clear how well they work in all cases. Therefore, broad bands may be desirable and the sensitivity to $\mathcal{C}_l$ may vary appreciably across the band. In such cases window functions, W_l^B, tell us the in-band sensitivity.

For experiments with small sky coverage, calculation of the elements of the Fisher matrix at every ℓ is not necessary. If the largest extent of the field is $\Delta\theta$ then $P_l(\cos\theta)$ and $P_{l+\delta l}(\cos\theta)$ are close to indistinguishable if $\delta l \lesssim \pi/\Delta\theta$. Thus one can choose a fine binning, enumerated by b, within each coarse band, B and assume that $F_{ll'}$ is constant for all l and l' within band b.

It some times may not be practical to calculate the Fisher matrix for individual multipole moments. It may be easier to parameterize the power spectrum in terms of these fine bins,

$$\mathcal{C}_l = \sum_b \chi_{b(l)} \mathcal{C}_b \tag{16}$$

and then calculate $F_{bb'}$ where

$$F_{bb'} = \mathrm{Tr}\left[(S+N)^{-1}\frac{\partial S}{\partial \mathcal{C}_b}(S+N)^{-1}\frac{\partial S}{\partial \mathcal{C}_{b'}}\right]. \tag{17}$$

Then one assumes that $F_{ll'} = F_{bb'}/(\delta l(b)\delta l(b'))$, where $\delta l(b)$ is the width in l of fine band, b. We divide by the widths because $\frac{\partial S}{\partial \mathcal{C}_l} \simeq \frac{\partial S}{\partial \mathcal{C}_b}/\delta l(b)$.

VI. OTHER SHAPES

It may be the case that an experiment reports a single measure of the power, but does so assuming a non-flat power spectrum shape. An historical example is the Gaussian auto-correlation function. Another possibility is that of an experiment measuring near the damping tail of the power spectrum, where assuming a flat shape may be a very bad approximation. Therefore we ask, if the amplitude of a non-flat power spectrum is estimated from the data, how can we calculate the theoretical predictions for this quantity?

To frame the question more precisely, we assume that Q is calculated from a data set via

$$Q = \frac{1}{2}F_{QQ}^{-1}\mathrm{Tr}\left[(\Delta\Delta - N)(S+N)^{-1}\sum_{l'}\frac{\partial S}{\partial Q}(S+N)^{-1}\right] \tag{18}$$

where the power spectrum is assumed to be of the form $Q\mathcal{C}_l^{\rm shape}$. Using similar manipulations as before, we find that the expectation value of Q, under the assumption that the true power spectrum is $\mathcal{C}_l$, to be

$$\langle Q \rangle = \frac{\sum_l \mathcal{C}_l \sum_{l'} F_{ll'} \mathcal{C}_{l'}^{\rm shape}}{\sum_{ll'} F_{ll'} \mathcal{C}_l^{\rm shape} \mathcal{C}_{l'}^{\rm shape}} = \sum_l \left(W_l^Q / l \right) \mathcal{C}_l . \quad (19)$$

Thus we have the prescription for comparing the estimated amplitude to the predicted amplitude.

Note that there is no clearly preferable means of converting the estimate of the amplitude, Q, into a measure of average power. One possible prescription is:

$$\mathcal{C}_Q = Q \sum_l \left(W_l^Q / l \right) \mathcal{C}_l^{\rm shape} / \sum_l \left(W_l^Q / l \right) . \quad (20)$$

Such a conversion is only useful for plotting purposes. The ambiguity in the choice of normalization does not disturb our ability to confront data with theory. As long as we know W_l^Q, and the shape assumed, $\mathcal{C}_l^{\rm shape}$, we can make the theoretical prediction for Q.

VII. DATA REPORTING RECIPE

To summarize, our prescription for reporting power spectrum constraints is as follows:
1) Parameterize the power spectrum via Eq. 13 for some choice of bands.
2) Find the $\mathcal{C}_B$ that maximize the likelihood function.
3) Calculate the curvature matrix for these bands, $\mathcal{F}_{BB'}$, and also the log-normal offsets x_B defined in [1].
4) Calculate the band-power window functions, W_l^B, from $F_{ll'}$ (which can be calculated either via Eq. 5 or Eq. 17).

Steps 1 through 3 have been spelled out in more detail in [1]. We have no general prescription for the best parameterization of the power spectrum to use for a given dataset (*e.g.*, how many bands to use and whether or not to assume a flat shape). We expect that assuming a flat spectrum across each band will be a reasonable choice in most situations. Whatever parameterization is chosen, the analyst should ensure that in addition to the estimate and its uncertainties, he or she also provide the means with which to convert a theoretical power spectrum into a prediction for that estimate.

VIII. DISCUSSION

As emphasized in [1], approximate methods for simultaneous analysis of all relevant CMB data are a practical necessity. The use of band-power window functions, instead of variance window functions will improve the validity of the commonly used method of radical compression to band-powers.

The persistence of the use of variance window functions as opposed to band-power window functions (without even acknowledgment that this is, at best, an approximation) is possibly attributable to the fact that maximum-likelihood estimates have a very complicated dependence on the data and, in fact, do not even have strictly well-defined band-power window functions. This conjecture is supported by the fact that analyses using quadratic estimators have not suffered from this confusion, while almost all of those using likelihood analysis have. It should also be noted that in analyses of galaxy redshift surveys, where quadratic estimators are generally used to estimate the matter power spectrum, $P(k)$, the correct form of the window function is generally used, *e.g.* [17].

As signal-to-noise rises, it becomes increasingly important to use the correct window function. This is because the power spectrum estimate becomes sensitive to more pairs of data points than just the diagonal ones—even the off-diagonal ones with very small signal matrix elements. One can see from Eq. (7) that in the limit that $N \to 0$, all pairs (normalized to their expected signal) get weighted equally. In this limit the much more numerous off-diagonal pairs are extremely important to the determination of the band-power. As a rough guide, one can compare the size of the largest off-diagonal terms to the noise level to determine how well the variance window function will approximate the band-power window function. Also note that the approximation is generally worse for map datasets than difference datasets due to the fact that differencing reduces the off-diagonal correlations.

Steps toward the proper definition of the band-power window function were taken in [11] where the diagonal elements of the window function matrix in the s/n basis were used as a means of determining the sensitivity of an experiment to the power spectrum. Working in the signal-to-noise eigenmode basis [2,16] reduces the correlations, and we have seen that in the absence of correlations, the band-power window function is the variance window function. A similar procedure was used to calculate the window functions for the band-powers determined from the QMAP maps [18].

As individual datasets become more powerful in their ability to determine $\mathcal{C}_l$, they will report constraints in very narrow bands, decreasing the need for window functions which describe the in-band sensitivity. However, a useful role for window functions may remain for quite some time at both the low l and high l extremes of a datasets' power spectrum sensitivity. At these extremes one will need broad bands in order to have small error bars, and thus one will wish to know the shape of the in-band sensitivity.

ACKNOWLEDGMENTS

I am grateful to J. R. Bond for useful conversations on this subject, to G. Wilson for supplying the $\partial S/\partial \mathcal{C}_b$ necessary to make the MSAM band-power window functions in Fig. 1, and to G. Wilson and an anonymous referee who both helped to clarify the text. I am supported by the DOE, NASA grant NAG5-7986 and NSF grant OPP-8920223.

---

[1] J.R. Bond, A.H. Jaffe & L. Knox, Astrophys. J., in press (astro-ph/9808264).
[2] J.R. Bond, *Astrophys. Lett. Comm.* **32**, 63 (1995).
[3] J.R. Bond, A.H. Jaffe & L. Knox, Phys. Rev. D **57**, 2117 (1998).
[4] M. Tegmark, *Phys. Rev.* **D55**, 5895 (1997).
[5] A.J.S. Hamilton Mon. Not. R. Astr. Society, **289**, 285 (1997) [astro-ph/9701008].
[6] G. Wilson, et al., Astrophys. J., in press (astro-ph/9902047).
[7] J.G. Bartlett, M. Douspis, A. Blanchard & M. Le Dour, preprint astro-ph/9903045.
[8] C.H. Lineweaver, astro-ph/9801029; C.H. Lineweaver, astro-ph/9805326; C.H. Lineweaver, & D. Barbosa, Astrophys. J. , **496**, 624 (1998); S. Hancock, G. Rocha, A.N. Lasenby & C.M. Gutierrez, Mon. Not. R. Astr. Society, **294**, L1 (1998); M. Tegmark, astro-ph/9809201; J.G. Bartlett, A. Blanchard, M. Douspis & M. Le Dour, astro-ph/9810318; C.H. Lineweaver, astro-ph/9810334; L.M. Griffiths, D. Barbosa & A.R. Liddle, astro-ph/9812125.
[9] J. R. Bond, *Cosmology and Large Scale Structure*, ed. R. Schaeffer, Elsevier Science Publishers, Netherlands (1996).
[10] M. White and M. Srednicki, *Astrophys. J.* **443**, 6 (1995).
[11] L. Knox, *Astrophys. J.* **480**, 72 (1997).
[12] M.J. Devlin et al., Astrophys. J. **430**, L1 (1994); E. Cheng et al., Astrophys. J. **488**, L59 (1997); E. Torbet et al., preprint astro-ph/9905100 (1999).
[13] U. Seljak, *Astrophys. J.* **506**, 64 (1998).
[14] P. Oh, D.N. Spergel & G. Hinshaw, *Astrophys. J.* **510**, 551 (1999).
[15] L. Knox, *Phys. Rev.* **D52** 4307 (1995).
[16] E.F. Bunn & M. White, *Ap. J.* **480**, 6 (1997).
[17] A.J.S. Hamilton Mon. Not. R. Astr. Society **289**, 295 (1997) [astro-ph/9701009].
[18] A. de Oliveira-Costa, M. J. Devlin, T. Herbig, A. D. Miller, C.B. Netterfield, L. Page & M. Tegmark, *Astrophys. J.* **509**, 77L (1998) [astro-ph/9808045].