INTERFEROMETRIC OBSERVATION OF COSMIC MICROWAVE BACKGROUND ANISOTROPIES

MARTIN WHITE,1 JOHN E. CARLSTROM,2 MARK DRAGOVA,2 AND WILLIAM L. HOLZAPFEL2

Received 1997 December 16; accepted 1998 October 23

ABSTRACT

We present a formalism for analyzing interferometric observations of the cosmic microwave background anisotropy and polarization. The formalism is based on the $\ell$-space expansion of the angular power spectrum favored in recent years. Explicit discussions of maximum likelihood analysis, power spectrum reconstruction, parameter estimation, imaging, and polarization are given. As an example, several calculations for the Degree Angular Scale Interferometer and Cosmic Background Interferometer experiments are presented.

Subject headings: cosmic microwave background — cosmology: theory — techniques: interferometric

1. INTRODUCTION

The cosmic microwave background (CMB) has become one of the premier tools for understanding early universe astrophysics, classical cosmology, and the formation of large-scale structure. It has a wealth of information about the origin and evolution of the universe encrypted in its signal. The frequency spectrum is that of a blackbody at 2.73 K, confirming the prediction of the standard hot-big bang model (see Fixsen et al. 1996; Nordberg & Smoot 1998; or the review of Smoot & Scott 1997). The small-angular scale power spectrum of the temperature and polarization anisotropies contain information on the cosmological parameters and the structure that existed at decoupling (see, e.g., Bennett, Turner, & White 1997).

The predictions for a wide range of cosmological models are now well understood and theoretically secure (Hu et al. 1998, hereafter HSWZ). Experimentally, a flurry of results have been reported in the last 5 yr. The study of the microwave sky has moved beyond the “detection phase” and into the “imaging phase,” with the next generation of experiments planning to provide detailed information about the shape of the angular power spectrum of the temperature anisotropies over a wide range of angular scales.

The advent of low-noise, broadband, millimeter wave amplifiers (Popieszalski 1993) has made interferometry a particularly attractive technique for detecting and imaging low-contrast emission such as anisotropy in the CMB. An interferometer directly measures the Fourier transform of the intensity distribution on the sky. By inverting the interferometer output, images of the sky are obtained that include angular scales determined by the size and spacing of the individual array elements. In this paper we discuss a formalism for interpreting CMB anisotropies as measured by interferometers and examine what two upcoming experiments, the Degree Angular Scale Interferometer (DASI) and the Cosmic Background Interferometer (CBI), may teach us about cosmology.

Several previous papers have dealt with the analysis of CMB data from interferometers (Martin & Partridge 1988; Subrahmanyan et al. 1993; Hobson, Lasenby, & Jones 1995; Hobson & Magueijo 1996). In this paper we extend the work to make explicit contact with the multipole space ($\ell$ space) methods now commonly adopted in analyzing single-dish switching experiments, including power spectrum estimation, parameter extraction, imaging, polarization, and mosaicking. We also give details of the upcoming DASI experiment.

The outline of the paper is as follows. We begin with a discussion of instruments past, present, and future in § 2. Foregrounds and point sources are discussed in § 3. The theoretical formalism for analyzing temperature anisotropies is outlined in § 4, while applications, including maximum likelihood estimation of parameters and power spectrum reconstruction, are treated in § 5. Increasing the sky coverage, and hence the resolution of the instrument in $\ell$ space, is introduced in § 6. Making images of the microwave sky is addressed in § 7. Polarization is treated in § 8. Finally, § 9 contains our summary and discussion.

2. INSTRUMENTS

The use of interferometers to study fluctuations in the CMB goes back over a decade (see Table 1). Early work using the VLA and Australia Telescope Compact Array (ATCA) concentrated on small angular scales, reporting a series of upper limits. Recently the Cambridge Anisotropy telescope (CAT; O’ Sullivan et al. 1995) has reported a detection of anisotropy on subdegree scales at low frequencies (15 GHz). Several groups are now planning to build interferometers that operate at higher frequencies and over a larger range of angular scales, with sensitivities that should enable them to map in detail the CMB anisotropy spectrum from $\ell \sim 10^2$ to $10^3$.

An interferometric system offers several desirable features: (1) It directly measures the power spectrum of the sky, in contrast to the differential or total power measurements. Images of the sky can then be created by aperture synthesis. (2) Interferometers are intrinsically stable, since only correlated signals are detected; difficult systematic problems that are inherent in total power and differential measurements are absent in a well-designed interferometer. This considerably reduces signals from ground pickup and near-field atmospheric emission. (3) They can be designed for continuous coverage of the CMB power spectrum with the angular spectral resolution determined by the number of fields imaged.

Motivated by these attributes, several groups are developing new interferometers targeted at measuring the
TABLE 1
CURRENT EXPERIMENTS TO MEASURE CMB TEMPERATURE ANISOTROPIES WITH INTERFEROMETERS

| Name   | Location   | $N_{\text{dish}}$ | Frequency (GHz) | Bandwidth (GHz) | Primary Beam | $\ell$ |
|--------|------------|-------------------|-----------------|-----------------|--------------|-------|
| OVRO   | US         | 6                 | 30              | 2.0             | 4'           | 6750  |
| VLA    | US         | 27                | 8               | 0.2             | 5'           | 6000  |
| Ryle   | England    | 8                 | 15              | 0.4             | 6'           | 4500  |
| BIMA   | US         | 10                | 30              | 0.8             | 6'           | 4300  |
| ATCA   | Australia  | 6                 | 9               | 0.1             | 8'           | 3400  |
| T-W    | US         | 2                 | 43              | ...             | 2'           | 20–100 |
| CAT    | England    | 3                 | 13–17           | 0.5             | 2'           | 339–722|
| VSA    | Canary Islands | 15          | 26–36           | 2.0             | 4'           | 130–1800 |
| DASI   | South Pole | 13                | 26–36           | 10.0            | 3'           | 125–700|
| CBI    | Chile      | 13                | 26–36           | 10.0            | 44           | 630–3500|

NOTE.—There are published upper limits from VLA, Ryle, and ATCA. The CAT has published a detection, while VSA, DASI, and CBI are expected to begin operations around 1999–2000. The location, number of dishes/horns, frequency, and (approximate) coverage in $\ell$ space are listed.

$^a$ Carlstrom, Joy, & Grego 1996.
$^b$ Fomalont et al. 1984, 1988, 1993; Knoke et al. 1984; Martin & Partridge 1988; Hogan & Partridge 1989; Partridge et al. 1997.
$^c$ Jones 1997.
$^d$ Cooray et al. 1997.
$^e$ Subrahmanyan et al. 1993, 1998.
$^f$ Timbie & Wilkinson 1990.
$^g$ O’Sullivan et al. 1995.
$^h$ See Jones 1996.
$^i$ See Halverson et al. 1998.
$^j$ http://astro.caltech.edu/~tpj/CBI/.

Anisotropy in the CMB on subdegree angular scales. Three instruments are currently under construction: the Very Small Array (VSA) in Cambridge, DASI, and CBI. DASI and CBI are parallel projects based at Chicago and Caltech. In this paper we will concentrate on the DASI instrument, although the formalism is completely general and applies equally to any interferometer.

DASI is an interferometer designed to measure anisotropies in the CMB over a large range of scales with high sensitivity. The array consists of 13 closely packed elements, each of 20 cm diameter, in a configuration that fills roughly half of the aperture area with a threefold symmetry (see Fig. 1). Each element of the array is a wide-angle corrugated horn with a collimating lens. DASI uses cooled

More information on DASI can be found at http://astro.uchicago.edu/dasi.

![Fig. 1](image_url)

**Fig. 1.**—(a) Physical configuration of the DASI receivers. The axes are graduated in wavelengths, and physical separations are obtained by multiplying by $\lambda = 1$ cm. Correlations are measured between all pairs of horns, so the instrument samples 78 baselines. Note the threefold symmetry of the array, which means that 26 different $u$'s are sampled. (b) The sensitivity of DASI vs. $\ell = 2\pi u$ for the configuration in (a). The window functions are shown for each of the 26 different $u$'s sampled, and the bold line traces the outline of these 26 window functions.
HEMT amplifiers running between 26 and 36 GHz with a noise temperature of less than 15 K. The signal is filtered into ten 1 GHz channels. DASI will operate at the South Pole.

3. FOREGROUNDS

In order to estimate the contribution of primordial fluctuations to the observed signal, it is necessary to estimate the contribution from foreground contaminants. Some reviews of the situation with regard to astrophysical foregrounds can be found in Brandt et al. (1994), Toffolatti et al. (1994, 1997), and Tegmark & Efstathiou (1996). We refer the reader to these papers for more details and lists of references. Below we summarize some of the more important points for the DASI experiment.

3.1. Atmosphere

The far field of a very compact array is actually quite near the instrument; for DASI this distance is only a few hundred meters. Thus everything beyond this distance, including the atmosphere, will be imaged by the interferometer. Taking a standard model for the static brightness of the atmosphere with a temperature $T$ and zenith opacity $n$, the brightness varies with zenith distance $\theta$ as

$$T \left[ 1 - \exp \left( \frac{-n}{\cos \theta} \right) \right]. \quad (1)$$

Its effect appears as a constant term plus a slope and a very slight curvature. Since the interferometer rejects low spatial frequencies, this atmospheric contribution is negligible in the final image.

The dynamic atmosphere causes a fluctuating brightness we must look through, and if it is in the far field it will be correlated and appear as excess noise (Church 1995). The Python 5 experiment observed from the South Pole during the austral summer of 1997 at a frequency of 45 GHz and covered angular scales comparable with those of DASI. From the level of atmospheric noise, we estimate that the atmosphere will contribute only about 10% to the total system noise (Lay et al. 1998).

3.2. Galactic Foregrounds

3.2.1. Synchrotron

Because of its low operating frequency, the main foreground DASI will need to contend with is Galactic synchrotron radiation produced by electrons spiraling in the Galactic magnetic field. The specific intensity of synchrotron emission roughly follows a power law in frequency $I_{\nu}^{\text{sync}} \propto \nu^b$ with spatially varying index and amplitude (Lawson et al. 1987; Banday & Wolfendale 1991; Platania et al. 1998). The index, $b_{\text{sync}}$, varies from $-0.1$ to $-1.3$, with a mean of $-0.8$. There is some evidence that the index steepens at higher frequencies. The angular power spectrum $C_{\ell} \sim \ell^{-3}$ for $\ell < 10^2$ (see eq. [6] for a definition of $C_{\ell}$).

3.2.2. Free-Free

Free-free emission, also known as bremsstrahlung, is due to scattering of unbound particles, typically electrons off nuclei, e.g., $ep \rightarrow ep\gamma$. The spectral index $\beta_{\text{ff}}$ depends on the temperature and density of the charged particles but is in the range $-0.13$ to $-0.16$ for typical electron density and temperature values for the interstellar medium (Bennett et al. 1992). At high Galactic latitudes, free-free emission is expected to dominate over synchrotron emission at around 40 GHz (Bennett et al. 1992). No direct maps of free-free emission exist, although there is a possible correlation between free-free emission and H$\alpha$ emission (Bennett et al. 1992). Since the H$\alpha$ maps contain striping, the significance of the correlation is not easy to assess. If there is a strong correlation, the free-free spectrum can be predicted from H$\alpha$ measurements at a Galactic latitude of $\sim 20^\circ$ (Reynolds 1992) plus fundamental physics to determine $\beta_{\text{ff}}$. There is evidence, however (Kogut et al. 1996; Leitch et al. 1997; de Oliveira-Costa et al. 1997; Kogut 1997), that free-free emission may be correlated with dust emission near the north celestial pole and that this “hot” ($10^5$–$10^6$ K) component may not emit H$\alpha$ (for an alternative explanation of the correlation in terms of spinning dust grains, see Draine & Lazarian 1998). A correlation between free-free emission and dust would imply $C_{\ell} \sim \ell^{-2.5-3}$ (Schlegel, Finkbeiner, & Davis 1998; Wright 1998).

3.3. Extragalactic Foregrounds

The fluctuations from extragalactic sources have been modeled by Toffolatti et al. (1994, 1997) and Franceschini et al. (1991). The source models are robust below 100 GHz but uncertain to almost an order of magnitude well above 100 GHz.

The angular dependence of uncorrelated point sources is of course that of white noise: $C_{\ell} \propto \ell^0$. There is some evidence that radio sources are correlated (Peacock & Nicholson 1991), but the non-Poisson contribution to the anisotropy is always smaller than the Poisson contribution (Toffolatti et al. 1994, 1997). Providing that sources exist over a large range of distances from us, any correlation in the sources at small scales is significantly diluted by projection.

We show in Figure 2 the angular power spectrum, $\ell(\ell+1)C_{\ell} \propto \ell^2$, associated with point sources at 30 GHz, assuming that we subtract all sources brighter than 30 mJy. We have taken the luminosity function of VLA FIRST (Becker, White, & Helfand 1996) radio sources at 1.5 GHz from Tegmark & Efstathiou (1996, eq. [43]) and extrapolated all the sources, assuming $I_\nu \propto \nu^\alpha$. For $\alpha > 0$ it is below the expected cosmological signal for the range of scales probed by DASI. Point-source subtraction for DASI will be facilitated by using the ATCA to map the DASI observing region at 16–26 and 43 GHz (the DASI observing region also overlaps with regions observed by Python and planned observations by Boomerang and Beast).

3.4. Foreground Subtraction

We will for definiteness here consider the DASI instrument, although our general conclusions will hold for the other planned instruments with similar frequency coverage. In the absence of external information about foreground emission, we can use the 10 GHz bandwidth of DASI to marginalize over the unknown amplitude of a foreground component. Since we are working at low frequency, the dominant foregrounds are synchrotron and free-free emission, which have similar spectral indices. Thus one is led to consider fitting out a single component. Using the formal-
The exponential of a phase factor. For each point source, the phase factor is the period of the wave. Assuming for now a monochromatic representation an average over a time long compared with the pointing to the same position on the sky, and angle brackets are the electric field vectors measured by two telescopes with a (1996, eq. [43]) and extrapolating 1996) radio sources at 1.5 GHz as provided by Tegmark & Efstathiou scales probed by DASI. The dashed and dotted lines are the angular power dimensionless units. The horizontal line represents the range of angular between the source and the two telescopes in units of the have been removed.

...known to be low in foregrounds, but the need for a large sky power spectra) can improve the separation of foreground spatial properties of the foregrounds (e.g., their DASI observations. As an example, assumptions about the parametric information (foreground maps) when interpreting the band of DASI and indicates that we would like to use external method. It may be possible to reduce this by roughly a factor of 2 (Tegmark 1997a; White 1998).

The increase in the error bar is due to the small operating band of DASI and indicates that we would like to use external information (foreground maps) when interpreting the DASI observations. As an example, assumptions about the spatial properties of the foregrounds (e.g., their \( \ell \)-space power spectra) can improve the separation of foreground and signal. The DASI will originally operate in a region known to be low in foregrounds, but the need for a large sky coverage will eventually require operation in regions with significant foregrounds. Because of the dependence on the region of sky surveyed, we will focus on the instrument characteristics from now on and not include the increase in the error bars expected from foreground subtraction.

4. FORMALISM

4.1. The Visibility

An interferometer measures \( \langle E_1 E_2^\ast \rangle \), where \( E_1 \) and \( E_2 \) are the electric field vectors measured by two telescopes pointing to the same position on the sky, and angle brackets represent an average over a time long compared with the period of the wave. Assuming for now a monochromatic source of radiation and working in the Fraunhofer limit, the average of the product of electric fields is the intensity times a phase factor. For each point source, the phase factor is the exponential of (i times) the geometric path difference between the source and the two telescopes in units of the wavelength. Taking the integral over the source/emitter plane gives the Fourier transform (FT) of the observed intensity (i.e., the sky intensity multiplied by the instrument beam).

The fundamental observable for the interferometer is thus a “visibility,” which is the FT of the sky intensity multiplied by the primary beam or aperture function (Tompson, Moran, & Swenson 1986):

\[
V(\mathbf{u}) \propto \int d\mathbf{\bar{x}} \ A(\mathbf{\bar{x}}) \Delta T(\mathbf{\bar{x}}) e^{2\pi i \mathbf{u} \cdot \mathbf{\bar{x}}}, \tag{2}
\]

where \( \Delta T \) is the temperature (fluctuation) on the sky, \( \mathbf{\bar{x}} \) is a unit three-vector, and \( \mathbf{u} \) is the conjugate variable, with dimensions of inverse angle measured in wavelengths. \( A(\mathbf{\bar{x}}) \) is the “primary” beam and is typically normalized to unity at peak, which is \( \sqrt{2\pi} \) larger than the usual normalization of a Gaussian beam. [By requiring \( A(0) = 1 \) we ensure that the area of the aperture in the \( \mathbf{u} \) plane is unity.] The spacing of the horns and the position of the beam on the sky determine which value of \( \mathbf{u} \) will be measured by a pair of antennae in any one integration. The size of the primary beam determines the amount of sky that is viewed and hence the size of the “map,” while the maximum spacing determines the resolution.

Typically, the field of view of the interferometer is small, so in what follows we will make a small-angle approximation and treat the sky as flat (see also the Appendix). This is a very good approximation for the upcoming experiments (CBI, DASI, VSA) and leads to significant simplification in the formalism. If the primary beam \( A(\mathbf{x}) \), or more generally the area of sky surveyed, is well localized, the integral is only over a very small range of \( \mathbf{\bar{x}} \). Denoting the center of the beam by \( \mathbf{x}_0 \), we can write \( \mathbf{\bar{x}} = \mathbf{x}_0 + \mathbf{x}_1 \) with \( x_1 \ll x_0 = 1 \) and \( x_1 \cdot x_0 = 0 \), then \( x_1 \) is a two-dimensional vector lying in the plane of the sky. We will denote two-dimensional vectors by boldface roman type and vectors in three dimensions or other spaces by boldface italic type. A vector name that is neither boldface roman nor boldface italic indicates the length of that vector, e.g., \( x_1 = |x_1| \).

Finally, we should remark on one subtlety in the statistical analysis of a “random” component like the CMB fluctuations. It is common to assume that the temperature fluctuations in the CMB are realizations of a random field, so \( T(\mathbf{x}) \) is a real random field on \( R^2 \) (see below). Since \( V(\mathbf{u}) \) is the FT of \( T(\mathbf{x}) \), it is also a random field; however, it is complex. Since \( T \) is real, it follows that \( V^\ast(\mathbf{u}) = V(-\mathbf{u}) \), so \( V(\mathbf{u}) \) is a complex random field with independent degrees of freedom only over the half-plane (i.e., twice as many as \( T \) over half the area). This restriction to the half-plane will be important when it comes to parameter estimation. An alternate formulation of the problem, which turns out to be equivalent, is to define a “real” visibility in terms of cosine (sine) transforms whenever, e.g., \( u_x > 0 \) (\( u_x < 0 \)). We will not give this parallel development here, as it is exactly equivalent to the complex case we will discuss.

4.2. The Sky Power Spectrum

To proceed, notice that in the small field-of-view approximation, the visibility is the convolution of the FT of the sky intensity (temperature) and the FT of the primary beam. Thus if we knew the power spectrum of the sky, we could find the power spectrum of the visibilities by convolution with the FT of the primary beam (see below). In this section we will concentrate on the sky power spectrum and neglect the effect of the primary beam, so we set \( A = 1 \) for now.
We usually assume that our theory has no preferred direction, i.e., it is rotationally invariant. In the flat-sky approximation, this rotational invariance becomes a translational invariance on the plane. This means that the (double) FT of the sky correlation function becomes diagonal in $u$, i.e., “conserves momentum.” (We discuss the flat-sky approximation further in the Appendix.) We will call the diagonal part the sky power spectrum $S(u) = S(u)$, not to be confused with a flux.

The ability to perform Fourier analysis on the “flat” sky and the replacement of rotational by translational invariance is the principle advantage of the flat sky approximation. These advantages are only obtained in the small-angle limit, regardless of how one chooses to map the angular into $x$, so the reader should beware of “improvements” to the flat-sky approximation except under very special circumstances.

We can write the FT of the correlation function, depending on $u$ and $w$, as

$$\int dx_1 dx_2 \, C(x_1 \cdot x_2) \exp [2\pi i u \cdot (x_1 - x_2)]$$

$$\times \exp [2\pi i (w - u) \cdot x_1] , \quad (3)$$

where $C(\cos \theta)$ is the (dimensionless) correlation function for the CMB temperature fluctuations, defined below. If we expand

$$e^{2\pi i u \cdot x} = J_0(2\pi u) + 2 \sum_{m=1}^{\infty} i^m J_m(2\pi u) \cos [m \arccos (\hat{u} \cdot \hat{x})]$$

and use the symmetry of the problem to do the angular integrals, we find (e.g., Subrahmanyan et al. 1993) that the diagonal part

$$S(u) \propto \int_0^2 \cos \omega C(\omega) J_0(2\pi u \omega) , \quad (5)$$

where $\omega = |x_1 - x_2| = 2 \sin (\theta/2)$ and $d(\cos \theta) = \omega d\omega$. In the flat-space limit we extend the upper limit of the $\omega$ integration to infinity. Expanding the correlation (or two-point) function for the CMB temperature fluctuations as a Legendre series,

$$C(x_1 \cdot x_2) \equiv \exp \left\{ \frac{\Delta T}{T} (x_1, \frac{\Delta T}{T} (x_2) \right\}$$

$$= \frac{1}{4\pi} \sum_{\ell} (2\ell + 1) C_\ell P_\ell(x_1 \cdot x_2) , \quad (6)$$

and using Gradsteyn & Ryzhik (1980, eq. 7.251[3]), we obtain

$$S(u) = \frac{1}{2\pi u} \sum_{\ell} (2\ell + 1) C_\ell J_{2\ell + 1}(4\pi u) . \quad (7)$$

In evaluating this sum, care must be taken for regions of the spectrum where $C_\ell$ is nearly scale invariant because of significant cancellations.

For large $\ell$, $J_{2\ell + 1}$ is a sharply peaked function. Thus the two-dimensional power spectrum $u^2 S(u) \sim \ell(\ell + 1)C_\ell |_{\ell = 2\pi u}$. Direct numerical evaluation of equation (7) or requiring the rms fluctuation at zero lag to be the same in $u$ space as $\ell$ space (Gradsteyn & Ryzhik 1980, eq. 6.511[1]) allows us to write

$$u^2 S(u) \approx \frac{\ell(\ell + 1)}{(2\pi)^2} C_\ell \left|_{\ell = 2\pi u} \right. \quad \text{for } u \gtrsim 10 . \quad (8)$$

This approximation works at the few percent level for standard cold dark matter (sCDM) when $u \gtrsim 10$ or $\ell \gtrsim 60$.

We show $S(u)$ versus $u$ for a selection of CDM models normalized to the COBE 4 yr data in Figure 3. The range of angular scales that will be probed by DASI and CBI are shown as the solid lines across the top of the figure. While these models were chosen primarily to fit the large-scale structure data, all of the models shown provide reasonable fits to the current CMB data. The parameters for the models are given in Table 2.

![Fig. 3a](image1.png)

**Fig. 3a**—Two-dimensional power spectrum per logarithmic interval in $u$ for a selection of COBE-normalized models chosen to provide good fits to the large-scale structure data (see text). The solid lines across the top of the plot illustrate the range of scales to which DASI and CBI will be sensitive. The left panel shows the expected uncertainties per bin for the sCDM model for 1 month of observing of 25 widely separated points on the sky. For display purposes we have placed the points on the sCDM curve. Each of the 26 baselines is shown, and the signal-to-noise ratio is about 1 in the highest $u$ bin. Although each baseline has independent receiver noise, the window functions for neighboring points overlap considerably, so the cosmological signal is very correlated between points. The right panel shows the result with 6 months of integration using a mosaicking strategy covering $\sim 400$ deg$^2$ of sky (again with $S/N \sim 1$ in the final bin). In this case, each of these points is completely independent. We have also shown the expected error bars for CBI, assuming a similar mosaicking strategy.
TABLE 2

| Name    | $\Omega_m$ | $\Omega_b h^2$ | $n$ |
|---------|------------|----------------|-----|
| sCDM... | 1          | 0.5            | 0.0125 | 1 |
| OCDM... | 0.5        | 0.6            | 0.0200 | 1 |
| tCDM... | 1          | 0.5            | 0.0250 | 0.8 |

Note.—All models have $\Omega_k = 0$ and have been normalized to the COBE DMR 4 yr data.

4.3. The Visibility Correlation Matrix

In theories that predict Gaussian temperature fluctuations the fundamental theoretical construct is the correlation matrix of the measured data. Since the data in our case are the visibilities measured at a set of points $u_i$, in what follows we will need to know the correlation matrices for the signal and noise in the various visibilities. The measured fluxes, $V(u)$, are

$$V(u) = \frac{\partial B_T}{\partial T} T_{\text{CMB}} \int dx \frac{\Delta T(x)}{T_{\text{CMB}}} A(x) e^{2\pi i xu} ,$$

(9)

where $\partial B_T/\partial T$ converts from temperature to intensity, $T_{\text{CMB}}$ is the CMB temperature, and $A(x)$ is the primary beam. The conversion factor from “temperature” to “intensity” is

$$\frac{\partial B_T}{\partial T} = \frac{2k_B}{c^2} \left( \frac{k_B T}{h} \right)^2 \frac{x^2 e^x}{(e^x - 1)^2} \approx \left( \frac{99.27 \text{ Jy sr}^{-1}}{\mu \text{ K}} \right) \frac{x^2 e^x}{(e^x - 1)^2} ,$$

(10)

where $B_T$ is the Planck function, $k_B$ is Boltzmann’s constant, $x \equiv h v/k_B T_{\text{CMB}} \approx v/56.84$ GHz is the “dimensionless frequency,” and 1 Jy equals $10^{-26}$ W m$^{-2}$ Hz$^{-1}$. In the Rayleigh-Jeans limit $\partial B_T/\partial T \approx 2k_B v/c$, where $v$ is the observing frequency. This is a good approximation for the frequencies of planned interferometers, the correction to the Rayleigh-Jeans assumption is 2% at 30 GHz, and we shall make it henceforth. We shall also set $c \equiv 1$, so $v = \lambda^{-1}$.

The FT of the primary beam\(^6\) is the autocorrelation of the FT of the point response, $g$, of the receiver to an electric field, $\bar{A}(u) = \bar{g} * \bar{g}(u)$, and

$$A(x) = \int du \bar{A}(u)e^{-2\pi i xu} ,$$

(11)

so using the fact that the power spectrum is diagonal in $u$, we have

$$C_{ij}^V \equiv \langle V^*(u_i)V(u_j) \rangle = (2k_B T_{\text{CMB}} v^2)^2 \int d^2w \bar{A}^*(u_i - w)\bar{A}(u_i - w)S(|w|) .$$

(12)

Notice that the visibilities are uncorrelated if $|u_i - u_j|$ is larger than (twice) the width of $\bar{A}$, which defines the bin size, $\Delta u$.

We are now in a position to define the window function $W_i(u)$, which, when convolved with the power spectrum, defines the visibility correlation matrix $C_{ij}^V$. From the above,

$$C_{ij}^V = (2k_B T_{\text{CMB}} v^2)^2 \int d^2w S(w)W_i(w) \text{;}$$

(13)

with

$$W_i(|w|) \equiv \int_{0}^{2\pi} d\theta_w \bar{A}^*(u_i - w)\bar{A}(u_i - w) .$$

(14)

Note that since $\bar{A}$ has compact support, the maximum of $W_i$ scales as $u_i^{-1}$ for $u_i \gg \Delta u$. Since the noise per visibility is independent of $u_i$, the signal-to-noise ratio drops as $u^{-2}$ for a scale-invariant spectrum, $S(u) \propto u^{-2}$. Also note that both $W_i$ and $S(u)$ are positive semidefinite, so the visibilities are never anticorrelated, unlike single-dish (chopping) experiments.

In general, the distribution of the electric field in the horn aperture is close to a pillbox times a Gaussian distribution. Because of the finite aperture, $\bar{A}$ has compact support. Usually the response is independent of $\hat{u}$; we will call the $|u|$ for which $\bar{A}$ vanishes $\Delta u$. In order to obtain a simple estimate of our window function, it is a reasonable first approximation to take $\bar{A}$ equal to the autocorrelation of a pillbox of radius $D/2$ where $D$ is the diameter of the dish. Specifically,

$$\bar{A}(u) = \frac{2\bar{A}_s}{\pi} \left( \arccos \frac{u}{D} - \frac{u\sqrt{D^2 - u^2}}{D^2} \right)$$

(15)

if $u \leq D$ and zero otherwise. Thus in this simple example $\Delta u = D$. If we require $A(0) = 1$, this must integrate to unit area, so $\bar{A}_s^{-2} = \pi(D/2)^2$ or the area of the dish.

In the case where all correlated signal is celestial, the correlation function of the noise is diagonal, with

$$C_{ij}^N = \left( \frac{2k_B T_{\text{sys}}}{\eta_A A_D} \right)^2 \frac{1}{\Delta u \tau_A n_b} \delta_{ij} .$$

(16)

Here $T_{\text{sys}}$ is the system noise temperature, $\eta_A$ is the aperture efficiency, $A_D$ is the physical area of a dish (not to be confused with $A(x)$), $n_b$ is the number of baselines\(^7\) corresponding to a given separation of antennae, $\Delta u$ is the bandwidth, and $\tau_A$ is the observing time. Typical values for DASI are

---

\(^6\) Throughout we will use a tilde to represent the FT of a quantity.

\(^7\) The number of baselines formed by $n_e$ receivers is $n_b = n_e(n_e - 1)/2$. 

---

Fig. 4.—Twenty-six diagonal elements of the correlation matrix $C_{ij}^V$ and $C_{ij}^N$ for COBE-normalized sCDM and a pillbox aperture function. The noise is assuming 1 day of operation of DASI and is per visibility. We have shown results for one pointing, without the instrument rotation necessary to fill the $\omega \nu$ plane. The signal in each of these visibilities will be highly correlated because of strongly overlapping window functions (see text).
$T_{\text{sys}} = 20 \text{ K}, \eta_A \sim 0.8$, dishes of diameter 20 cm, $n_b = 3$, and $\Delta T = 10 \text{ GHz}$ (in $10 \times 1 \text{ GHz}$ channels). For CBI the dishes are 5 times larger, with the other numbers about the same. We show in Figure 4 the diagonal entries of $C^V$ and $C^N$ for one pointing and 1 day of observing with DASI.

Using equations (12) and (16), we can provide a rough estimate of the signal-to-noise ratio expected in a given visibility. For a given by equation (15), we have

$$\frac{C^V}{C^N} \sim \left( \frac{T_{\text{CMB}}}{T_{\text{sys}}} \right)^2 \Delta t \tau_n \left( \frac{\Delta u}{u} \right)^2 u^2 \delta(u) ,$$

which, given the numbers above, yields $C^V / C^N \sim 10^3 (\Delta u / u)^2$ for 1 day of integration, assuming a COBE-normalized, scale-invariant spectrum.

5. COMPARISON WITH THEORY

5.1. Likelihood Function

In stochastic theories $\Delta T(\hat{\mathbf{x}})$ is a (Gaussian) random variable with zero mean and dispersion given by the $C$. Thus the visibilities measured by the interferometer will be Gaussian random variables with zero mean and dispersion $C^V + C^N$. For a given set of measured visibilities one can test any theory, or set of $\mu_i$, by constructing the likelihood function (for complex variables $V$)

$$L(C) = \frac{1}{\pi \text{det } C} \exp \left[ -V^*(\mathbf{u}) C_{ij}^{-1} V(\mathbf{u}) \right] ,$$

where $C_{ij} = C^V_{ij} + C^N_{ij}$ is the correlation matrix of visibilities at $\mathbf{u}$ and $\mathbf{u}_i$ (Hobson et al. 1995). Note that the visibilities are complex, so the likelihood function is slightly different than for the case of real Gaussian random variables, e.g., the “missing” factor of $\frac{1}{2}$ in the exponent. The restriction to the half-plane, however, ensures that the number of degrees of freedom is the same as for the real case (see Hobson et al. 1995).

Given a set of data $\{ V_i \}$, one can proceed to test theories using the likelihood function. Confidence intervals for parameters and relative likelihood of theories are calculated in the usual way.

5.2. Power Spectrum Estimation

There are several ways one could consider estimating the angular power spectrum from a set of visibilities. Conceptually, the simplest is to average $|V(\mathbf{u})|^2$ in shells of constant $|\mathbf{u}|$. This gives a (noise-biased) estimate $C$ at $\ell = 2\pi u$, convolved with the window function $W_{\ell}$.

A more sophisticated method is to define as a “theory” a set of bandpowers, i.e., define the power spectrum $\ell (\ell + 1)C_{\ell}$ as a piecewise constant, in $N_{\text{bands}}$ bands. One then maximizes the likelihood function for this “theory”; the result is the best-fitting power spectrum, binned into groups at similar $\ell$. Such an approach has been used on the COBE 4 yr data by Gorski et al. (1996) and Bunn & White (1996) and discussed extensively by Bond, Jaffe, & Knox (1998). It is the method we will advocate here.

As was pointed out by Bond et al. (1998), it is particularly simple to find the maximum of the likelihood function for such a “theory” using a quadratic estimator (see also Tegmark 1997b) that takes as an input a trial theory. In this case, the theory is a set of “band powers” chosen to cover the $\ell$ range of interest and be approximately independent. Iteration of the quadratic estimator is equivalent to Newton’s method for finding the root of $dL / dp$, where $p$ is a parameter on which $L$ depends.

Specifically, from an estimate of the band powers, $u^2 S(u) = \hat{p}_g$ for $u \in B_g$ (where $g = 1, \ldots, N_{\text{bands}}$); an improved estimate is $p_g = \hat{p}_g + \delta p_g$, with

$$\delta p_g = (\text{tr}[\hat{C}^{-1} \hat{C}_g \hat{C}_g^{-1} \hat{C}_g^{-1}])^{-1} \text{tr}[ (C - \hat{C}) (\hat{C}_g^{-1} \hat{C}_g^{-1} )]$$

(Bond et al. 1998). Here $\hat{C}$ indicates the (theoretical) correlation matrix evaluated at the initial estimate $\hat{p}_g$, with $C_g \equiv dC / dp_g$, and $C \equiv V^*(\mathbf{u}) V(\mathbf{u})$ indicates the matrix formed by the data. Iteration of this procedure (e.g., from an initially flat spectrum $p_g = \text{constant}$) converges to the maximum likelihood estimate of the power spectrum.

Since the parameters $p_g$ are chosen to be the (constant) values of $u^2 S(u)$ across the band $B_g$, and the noise is assumed to be independent of the level of cosmological signal,

$$\hat{C}_{ij} = \left( 2 k_B T_{\text{CMB}} v^2 \right)^2 \int \frac{dw}{w^2} W_{ij}(w) .$$

We may gain some intuition for this expression by considering the simple case of one band and uncorrelated visibilities. Assume additionally that $W_{ij}(u)$ is independent of $u$. Then if we write $C = \sigma_i^2 \delta_{ij}$ and $\hat{C}_{ij} = \sigma_i \delta_{ij}$, we have

$$\delta p_g = \left( \sum_i \sigma_i^4 \right)^{-1} \sum_i \left| V_j \right|^2 \left| w_j \right| - \sigma_j^2 .$$

This becomes even simpler if all of the visibilities are at fixed $|\mathbf{u}|$. Then $\sigma^2$ is independent of $i$, and, if we assume the noise is as well, $\delta p$ becomes zero when

$$(\sigma^2)^2 = \frac{1}{N} \sum_{j=1}^N \left| V_j \right|^2 \left| w_j \right| - (\sigma^2)^2 ,$$

which is reminiscent of our simplistic estimate described at the beginning of this section.

5.3. CDM Parameter Estimation

It has become common (Scott & White 1995; Jungman et al. 1996; Bond, Efstathiou, & Tegmark 1997; Zaldarriaga, Spergel, & Seljak 1997; Stompor & Efstathiou 1998) to ask how well we could measure theory parameters “on average” given a set of measurements $\{ V_i \}$ that are “typical.” Imagine that our theory is defined by a set of parameters $\{ p_g \}$ and that the sky corresponds to this theory with values of parameters $\hat{p}_g$. In this case,

$$\langle - \ln L \rangle = \text{tr}[\hat{C}^{-1} \ln C] ,$$

where $\hat{C}_{ij} = \langle V_i^* V_j \rangle$ denotes $C_{ij}(\hat{p})$. The precision to which we can measure $\langle p_g \rangle$, an input theory $\hat{C}$, is given by the second derivative matrix of $\langle - \ln L \rangle$:

$$\frac{\partial^2 \langle - \ln L \rangle}{\partial p_g \partial p_g} = \text{tr}[\hat{C}^{-1} \hat{C}_{,g} \hat{C}_{,g}^{-1} \hat{C}_{,g}] ,$$

with $C_{,g} \equiv dC / dp_g$. Thus

$$\langle p_g - \langle p_g \rangle \rangle = \{ \text{tr}[\hat{C}^{-1} \hat{C}_{,g} \hat{C}_{,g}^{-1} \hat{C}_{,g}] \}^{-1} ,$$

where $\langle \cdots \rangle$ denotes an average with respect to the likelihood function $L$.

As an example of how well planned interferometers will constrain cosmological models, we can ask how well DASI would be able to determine the cosmological parameters
We show in Table 3 the relative uncertainties on $p_\kappa$ with a “prior” 20% uncertainty in $C_{10}$. (We do not include the increase in the errors from foreground subtraction here.) The number of bins is fixed at $N_{\text{bin}} = 20$, or $10^3$ deg$^2$ of sky coverage, for simplicity. For 3 months of observing an sCDM sky, DASI would be system noise–limited for $u \gtrsim 50$. Note that there is a strong dependence of the estimated errors on the input theory (and the parameter set chosen). We have chosen the three theories here to explore this dependence. For the open model (which has $\Omega_\Lambda = 0.5$), the peaks in the power spectrum are at higher $u$ than the critical density models (see Fig. 3). This means that there is less information (from the higher peaks) available to break the degeneracy in parameter variations, leading to larger uncertainties when the other parameters are integrated out. To increase the precision with which cosmological parameters could be determined in an open model, one would need to extend the coverage to higher $|u|$ using CBI. If this is done, the angular scale of the features in the open model is better extended to higher $u$. The tilted model, the signal-to-noise ratio is lower at large $u$, which accounts for the slightly larger uncertainties on, e.g., the spectral slope $n$. However, the higher baryon fraction leads to greater sensitivity to $\Omega_b$, and the variations with $\Omega_\text{mat}$ are less correlated with $\Omega_b$.

For all of these theories the uncertainties are larger because of the large correlations between parameters: if the bin size or noise is too large, one cannot distinguish between variations in different parameters, inflating the marginalized errors. To decrease the correlations one must increase the resolution in $u$, which means obtaining more sky coverage, or increase the total range of $u$ covered, which means combining DASI and CBI. The combination of DASI and CBI is particularly powerful, with the marginalized errors for each model being below 20% for all parameters within 1 yr of observing.

### 6. MOSAICKING

#### 6.1. Increasing Resolution

The resolution we have in $u$ space is limited by the amount of sky that we have surveyed, which for a single pointing is equal to the size of the primary beam. By combining several contiguous pointings of the telescope we can increase the amount of surveyed sky and therefore increase the resolution in $u$ space. This is known as mosaicking. The idea here is to measure the visibility as a function of posi-

| Name           | 3 Months | 6 Months | 12 Months |
|----------------|----------|----------|-----------|
|                | $\Omega_\text{mat}$ | $h$ | $\Omega_b$ | $n$ | $\Omega_\text{mat}$ | $h$ | $\Omega_b$ | $n$ | $\Omega_\text{mat}$ | $h$ | $\Omega_b$ | $n$ |
| sCDM ........  | 21        | 40       | 50        | 9  | 15        | 26    | 34    | 7  | 12        | 19    | 27    | 5  |
| OCDM .......... | ...       | ...      | ...       | 62 | 73        | ...   | 45    | ... | 53        | ...   | 91    | 33 |
| tCDM .......... | 30        | 71       | 40        | 16 | 17        | 39    | 25    | 10 | 11        | 23    | 17    | 7  |

Note.—The percentage relative uncertainties for the cosmological parameters assuming 3, 6, or 12 months of integration of the DASI over $10^3$ deg$^2$ ($N_{\text{bin}} = 20$) for the theories discussed in the text. We have included a “prior” corresponding to 20% uncertainty in $C_{10}$. For the open models several parameters are very correlated, because DASI alone does not probe to high enough $u$. For sCDM with $N_{\text{bin}} = 20$ and $T_{\text{sys}} = 20$ K, we are noise-limited for $u \gtrsim 50$ with 3 months observing time. Once DASI and CBI are combined, the uncertainties on all parameters for all models are less than 20% for 1 yr of data.
tion \( y \) from the map or phase center,
\[
V_y(u) \propto \int d^2x \, A(x - y)T(x)e^{2\pi i u \cdot x}.
\]  

We then sample \( V_y \) at a series of points by repointing the entire telescope. Let us denote this by a sampling function \( \Pi(y) \), which will be a sum of delta functions. We compute the FT of \( \Pi(y)V_y \), which is simply the convolution \( \{\hat{\Pi} \ast \hat{V}\}(v) \), where
\[
\hat{V}(u, v) = \hat{\Pi}(v)\hat{T}(u + v).
\]

Here \( u \) is the original vector in the \( u-v \) plane, and \( v \) is the variable conjugate to \( y \).

If our sampling function is a sum of delta functions, then \( \hat{\Pi} \) is a sum of plane waves. Thus simply summing our \( V_y \) at points \( y_j \), with weight \( c_j \), changes our aperture in \( u \) space from \( \hat{A} \) to
\[
\sum_j c_j e^{-2\pi i v \cdot y_j} \hat{A}(v),
\]
which can be made much narrower. This is completely analogous to the usual case of Fraunhofer diffraction through many holes, which is treated in most textbooks on optics. The simplest example is when \( c_j = 1 \) and the \( y_j \) lie on a regular \( N \times N \) grid with spacing \( \delta \); then
\[
\sum_j e^{-2\pi i v \cdot y_j} = \exp \left[ -2\pi i v \cdot \delta \left( \frac{N - 1}{2} \right) \right]
\times \frac{\sin(N2\pi v_x \delta_x)}{\sin(2\pi v_x \delta_x)} \frac{\sin(N2\pi v_y \delta_y)}{\sin(2\pi v_y \delta_y)}.
\]

The second diffraction spike occurs at \( v_{x,y} = \delta_{x,y}^{-1} \). If we choose \( \delta_{x,y} \) to be sufficiently small (Nyquist sampling), this will be outside the range where \( \hat{A} \) vanishes.\(^8\) Figure 5 shows the gain for a \( 3 \times 3 \) mosaic.

\(^8\) If we sample at precisely the spatial Nyquist rate, gain variations from star to star will also alias power outside of the spatial frequency window.

In the absence of a preferred direction in the theory, the power spectrum of the sky is symmetric in \( \hat{u} \), so choosing a mosaicking be maximally rotationally symmetric within the observing constraints. In Figure 5 we compare the \( 3 \times 3 \) square with seven points laid out at the center and vertices of a regular hexagon, with all points equidistant from their neighbors.

Note that mosaicking does not increase the range of \( u \) to which we are sensitive (see eq. [28]); it simply enhances our resolution by allowing us to follow more periods of a given wave. Thus we retain our ability to reject long-period noise or foregrounds.

If the goal were simply a measurement of the power spectrum, it is just as efficient to use smaller telescopes (with intrinsically better \( u \) resolution) and integrate for longer as it is to use large dishes and mosaic to increase the resolution. The advantage of the later method is that it allows better imaging of each piece of sky for checks of systematics, non-Gaussian features, and foregrounds, and it is easier to avoid "dropouts" in the \( u \) coverage for the power spectrum.

There are two routes to analyzing mosaicked data. The first is to treat the visibilities \( V_y(u) \) as separate data, highly correlated in a calculable way, with apparently low resolution but much information in the correlations between visibilities. The theory correlation matrix of mosaicked data can be written in the form of equation (13) with a modified window function that allows different pointing centers for each visibility:
\[
W_{ij}(|w|) = \int d\theta_{w} \, \hat{A}^{*}(u_i - w)\hat{A}(u_j - w) \times e^{2\pi i (u_i - 3/2 \delta)(w - 3/2 \delta)}.
\]

We discuss this further in § 6.2 below.

The other route is to statistically weight the \( V(u) \) from the different \( y_j \) to form a synthesized data set with fewer visibilities and correlations and intrinsically higher resolution. This method is perhaps simpler to understand, but the weighting of the different \( y_j \) will probably not be

---

**Fig. 5a**

Aperture function \(|A(u)|^2 \) for (a) \( 3 \times 3 \) square mosaic with spacing \( D/2 \) and (b) seven points laid out in three rows (at the center and vertices of a regular hexagon) with all neighbors separated by \( D/2 \). Contours mark 1, 2, 3, and 4 \( \sigma \) from the peak, with 3 \( \sigma \) shown thick (1 \( \sigma = e^{-1/2} \) of peak power). The dashed lines indicate the unmosaicked aperture of the primary beam, the solid lines the result after mosaicking.
optimal for parameter estimation or power spectrum estimation.

6.2. Optimal Subspace Filtering

A set of data from several (mosaicked) pointings involves many measured visibilities, very few of which are independent. Thus likelihood analysis requires repeated inversion of a large matrix, which can be computationally quite slow. One method for increasing the efficiency of the calculation is to work with a subset of the data that contains most of the signal but has fewer elements. This transformation can be accomplished using an “optimal subspace filter.”

Consider first the case of a single (unmosaicked) pointing. For DASI the visibilities will be measured at a specified set of \( u \) at any one time, then the instrument will be rotated about an axis through the center of the aperture plane to obtain a different set of \( u \) with the same lengths but different orientations. In this way most of the \( u-v \) plane will be covered, allowing for imaging. The number of rotations will be chosen to almost fully sample the \( u-v \) plane at the highest resolution (largest baseline or largest \( u \)). This means that considerable oversampling of the smaller baselines must result, since the aperture \( A \) has a fixed width of \( Au \). Phrased another way, the correlation matrix \( C^f \) of our visibilities has the dimension \( \sim (u_{\text{max}}/Au)^2 \). However, only a fraction of the entries are uncorrelated and contain “new” information about the theory. The rest are primarily a measure of the noise in the experiment. Thus if we could perform a change of basis to those combinations that measure primarily the signal and those that measure primarily noise, we could work in a much smaller subspace of the data (the “signal” subspace) with little loss of information about the theory. In calculating the likelihood function, we need to invert \( C^f \). Since matrix inversion is an \( N^3 \) process, reducing the amount of data can dramatically decrease the processing time.

As a simple example of this technique, imagine that we measure a visibility twice, with independent noise in each measurement. The signal in these two measurements is totally correlated. The sum of the two measurements is primarily sensitive to the signal (with noise \( 1/\sqrt{2} \) of each individual measurement), while the difference measures primarily noise. Thus if we “change basis” to the sum and difference of the two measurements, we could work with the sum only with little loss of information. This would halve the size of the matrix we would need to invert in the likelihood function and speed processing by a factor of \( \sim 8 \).

Optimal subspace filtering (also known as the signal-to-noise eigenmode analysis or Karhunen-Loeve transform) is a method designed to estimate the best change of basis and subspace in which to work. Recent applications of this method to the CMB and large-scale structure can be found in Bond (1995), Bunn & Sugiyama (1995), Bunn, Scott, & White (1995), Bunn (1995), White & Bunn (1995b), Bunn & White (1996), Vogelezang & Szalay (1996), Bond & Jaffe (1996), Tegmark, Taylor, & Heavens (1997), and Tegmark et al. (1998). In the above example, \( C^f_{ij} = 1 \) for all \( i, j \). This matrix has two values \( 2 \) and \( 0 \) (orthonormal) eigenvectors \( (1, 1)/\sqrt{2} \) and \((1, -1)/\sqrt{2}\), respectively. The eigenvalues measure the signal-to-noise ratio carried by the combination of data points \( \mathbf{d} \cdot \mathbf{\psi}_n \), where \( \mathbf{\psi}_n \) are the (orthonormal) eigenvectors.

In general, it can be shown that to find the optimal subspace on which to project, one finds the eigenvalues and eigenvectors of the matrix \( \mathbf{C}^{-1/2} \mathbf{C} \mathbf{C}^{-1/2} \), where \( \mathbf{C} \) is the noise matrix in measurement \( i \). Since \( \mathbf{C} \) is positive definite and symmetric finding, the eigenvalues and eigenvectors are straightforward (e.g., by Jacobi transforms). Arrange the eigenvectors in decreasing order of eigenvalue. The first \( M \) eigenvectors in the sequence define the best \( M \)-dimensional subspace to use in filtering the noise from the data. Once increasing the dimension of the subspace adds eigenvectors whose eigenvalues are \( \ll 1 \), the gain for the extra computing burden is marginal.

The advantage of optimal subspace filtering when fitting mosaicked data is obvious. One wishes to cover the \( u-v \) plane fully with the small effective apertures \( A \) obtained after summing the individual pointings. Thus each pointing oversamples the \( u-v \) plane considerably in terms of its larger “unmosaicked” \( A \), and these visibilities are highly correlated. The optimal subspace filter identifies which combinations of the visibilities measure primarily the cosmological signal and which are mostly noise, allowing an optimal weighting to be given to the individual visibility sets while retaining the processing time advantage of a smaller data set.

7. IMAGING THE MICROWAVE SKY

In addition to power spectrum estimation and parameter constraints, one goal of DASI and CBI is to image the microwave sky. Since interferometers do not measure the temperature of the CMB sky directly, it is necessary to “invert” the measured visibilities to form a sky image. In general this inversion is not unique (the DC level, for example, remains unconstrained) and so must be regularized. On a field-by-field basis, images of the sky can be made using the usual methods of synthesis imaging in radio astronomy (e.g., Cornwell 1989) and will form a useful data-checking tool.

It is possible, however, to try to make a larger scale image of the microwave sky using other (statistical) techniques. On degree angular scales such as those probed by DASI, regularized images have been made from the ACME/Max (White & Bunn 1995a), Saskatoon (Tegmark et al. 1997), and CAT (Jones 1996) experiments using a variety of techniques. Perhaps the most straightforward procedure for going from the visibility data to a sky image is Wiener filtering. Wiener filtering has been used extensively on CMB and large-scale structure data in the past (e.g., Lahav et al. 1994; Bunn et al. 1994; Zaroubi et al. 1995; Bunn, Hoffmann, & Silk 1996), and the reader is referred to those papers for a description of the underlying theory. The implementation of the Wiener filter in the basis defined in § 6.2 is treated in many of the references given in that section.

The Wiener filter provides an estimate of the sky temperature at a point as a linear combination of the visibility data. If we assume that foregrounds and point sources have been removed from the visibility data and that the underlying sky is Gaussian, then Wiener filtering is the “optimal” imaging method. Because it is linear, under the assumption of a Gaussian sky, the error matrix in the map is also easy to calculate.

\[ \mathbf{C}^{-1/2} = N^{-1/2} \mathbf{R} \mathbf{C}^{-1/2}, \]

where \( N^{-1/2} \) is any square root of the noise correlation matrix: for example, that defined by Cholesky decomposition.
To proceed, let us imagine pixelizing the sky into very many small pixels at position \( x_n \), \( n = 1 \cdots N_{\text{pix}} \). This is purely a bookkeeping device that allows us to use a matrix notation for the Wiener filter; the continuum limit is obtained trivially with a sum over pixels replaced by an integral. We denote the underlying temperature fluctuation on the sky by \( t_s \), which is related to the \( j \)th visibility by \( v_j = w_{j\mu} t_s + n_j \), where the weight matrix \( w_{j\mu} \) is (see eq. [2])

\[
W_{j\mu} = e^{2\pi i u \cdot x_j} A_j(x_a) \tag{32}
\]

and \( n_j \) is the noise in the \( j \)th visibility. We have labeled \( A_j(x) \) with a subscript \( j \) to allow for the possibility of mosaicking.

On an algorithmic note, it is computationally simpler to split the visibility into its real and imaginary parts and modify \( W_{j\mu} \) to have cosine and sine components so that one deals only with real variables. The number of data points is then doubled, and the noise covariance matrix must be modified by a factor of 2 as well. We will use a complex notation, with the understanding that the implementation of the algorithm may be in terms of real valued data.

If we assume that the correlation matrices of the theory and noise are known (or computed from the data, see Seljak 1997b), then the Wiener-filtered estimate of \( t_s \) is

\[
t_s^W = C_{s\mu} W_{j\mu} C_{\nu\nu} W_{k\nu} + C_{\mu\nu}^{-1} V_k, \tag{33}
\]

where \( C_s \) and \( C_N \) are the signal (eq. [6]) and noise (eq. [16]) correlation matrices. The combination \( W C_s W^T \) is essentially \( C_v \) of equation (12); we have written it in this way to allow the possibility of including other data sets, as is discussed below. Notice that the form of the Wiener filter is “signal/(signal + noise),” as is usually the case.

An expression similar to equation (33) can be obtained for the (normal) error correlation matrix of the estimates (see the above references). The most serious problem is that the Wiener-filtered image constructed in this manner is missing large-scale power. This can be included by fitting to another data set that retains the long-wavelength modes at the same time as the visibility data. In this case one extends the data vector \( v_j \) to include the extra temperature data, with the associated \( W_{j\mu} \). In the case of a mapping experiment, \( W_{j\mu} \) is simply the beam. For most current degree-scale experiments, it is a beam modulated on the sky by a chopping pattern. We will defer a detailed analysis of how well we can reconstruct the sky under various observing strategies, and the noise properties of the images, to a future publication, where we also plan to discuss the effect of sky curvature.

8. POLARIZATION

Neither DASI or CBI will be polarization sensitive initially; however, some instruments operating at smaller angular scales (e.g., the VLA) are already sensitive to linear polarization. The (linear) polarization of the CMB anisotropies is an important theoretical prediction and can encode a great deal of information about the model (for a recent review see Hu & White 1997a, so we discuss it briefly here. Our analysis is very similar to the small-angle formalism developed in Seljak (1997a), although we caution the reader that we have a different sign convention for the \( E \)- and \( B \)-modes; see, e.g., HSWZ.

We consider here only the small-scale limit, so we treat the sky as a plane with a right-handed coordinate system on it so that the sky is the \( x-y \) plane. We define polarization in the horizontal (\( \hat{x} \)) and vertical (\( \hat{y} \)) directions to be \( Q > 0 \) and \( Q < 0 \), respectively. Polarization in the \( \hat{x} + \hat{y} \) and \( \hat{x} - \hat{y} \) directions is defined to be \( U > 0 \) and \( U < 0 \), respectively. In terms of light traveling to us along \( \hat{z} \), the intensity tensor \( I_{ij} \), with \( i, j = x, y \), is

\[
I_{ij} \propto \langle E_i^* E_j \rangle \propto T + Q \sigma_3 + U \sigma_1, \tag{34}
\]

where the angle brackets indicate an average over a time long compared with the frequency of the wave, \( E_i \) is the electric field component, and \( \sigma_3 \) are the Pauli matrices. We have neglected \( V \sigma_2 \), since this corresponds to circular polarization, which is not generated cosmologically. Under a rotation by an angle \( \psi \) around \( \hat{z} \), the temperature is clearly left invariant, while \( Q \pm iU \) transforms as a spin-2 tensor:

\[
(Q \pm iU) \rightarrow e^{\mp 2i\psi}(Q \pm iU). \tag{35}
\]

We expand \( Q \pm iU \) in basis states known as spin-spherical harmonics \( \pm 2 Y_{lm} \) with coefficients \( a_{2m} \), in analogy with the temperature fluctuations (Seljak & Zaldarriaga 1997; Kamionkowski, Kosowsky, & Stebbins 1997; Hu & White 1997b). We can form states of definite parity, \( a_{2m} \) and \( a_{2m} \), which are called \( E \)- and \( B \)-mode polarization, respectively (not to be confused with the \( E \)- and \( B \)-modes of the radiation), and have angular power spectra \( C_E \) and \( C_B \), like the temperature. In addition, the \( E \)-mode of the polarization is correlated with the temperature, so there is a fourth power spectrum: \( C_{EB} \).

The spin-2 spherical harmonics are better known in the context of quantum mechanics as Wigner functions:

\[
y_{lm} \propto \mathcal{D}_{-s,m}(\ell, \theta, \phi) = \langle \ell, s | R \ell m \rangle , \tag{36}
\]

where \( R = \exp[\text{ix} \cdot \mathbf{J}] \) is the rotation operator and \( J \) is the angular momentum operators. The case \( s = 0 \) is the well-known spherical harmonic. Thus we can relate \( \pm 2 Y_{lm} \) to the usual \( y_{lm} \) through raising and lowering operators (Zaldarriaga & Seljak 1997). The raising and lowering operators are differential operators, since differentiating \( \mathcal{D} \) with respect to its arguments inserts \( j_{+} \) or \( j_{-} + i j_{0} \) inside the bracket in equation (36). In the flat space case the operators are trivially \( \partial_{\alpha} \pm i \partial_{\beta} \), with an overall normalization of \( \ell^{-1} \) from the normalization of the angular momentum states.

Going to the flat-space limit (see the Appendix) and acting on the Fourier integral with \( \mathcal{D}_{-s,m} \), we find that the Fourier coefficients \( E_{m} \) and \( B_{m} \) are related to the measured \( Q \) and \( U \) through

\[
-\langle Q \rangle = \int d^2u E_{m} \cos 2\theta_{m} - B_{m} \sin 2\theta_{m} e^{2\pi i u \cdot x},
\]

\[
-\langle U \rangle = \int d^2u E_{m} \sin 2\theta_{m} + B_{m} \cos 2\theta_{m} e^{2\pi i u \cdot x}, \tag{37}
\]

where the signs reflect the definition of \( E \) and \( B \) in (HSWZ) and come from differentiating an imaginary exponential twice. (We caution the reader that the power spectra output by the program CMBFAST differ in the sign of \( C_{EB} \) from this convention; HSWZ. Obviously \( C_{EE} \) and \( C_{BB} \) are invariant under the sign change.) If we recall that changing the sign of the polarization rotates the polarization “vector” through 90°, we see that the \( E \)-mode has polarization always tangential in the \( u-v \) plane (Seljak 1997a; it would be radial under a sign change).

We now have four nonvanishing power spectra, all diagonal in \( u \), whose diagonal parts we can denote by \( S_T, S_E, S_{EB} \), and...
where we have replaced the sum over \( \text{Gradshteyn & Ryzhik (1980, eq. 8722[2])} \), we have

\[
C_{ll}^{T} \propto \int d^{2}u \bar{A}_{l}(\mathbf{u}) \bar{A}_{l}(\mathbf{u}) \sum_{x,y} a_{x}^{T} a_{y}^{Y} \bar{S}_{x}^{y}(\mathbf{u}) ,
\]

where \( X = (T, Q, U) \) and \( x, y = T, E, B \), with \( a^{TT} = 1 \), \( a^{QE} = -\cos 2\theta \), \( a^{QB} = \sin 2\theta \), \( a^{EE} = -\sin 2\theta \), and \( a^{UB} = -\cos 2\theta \).

Given a set of measurements, one can reconstruct the power spectrum of the temperature or the polarization, as is described in § 5.2. Alternatively, one can provide upper limits (or measurements) of a bandpower for polarization or temperature by calculating \( \mathcal{L}(Q^{2}) \), where \( Q^{2} \) is the amplitude of a flat power spectrum (i.e., \( \ell(\ell + 1)C_{\ell} = \text{constant} \)) for the TT, TE, EE, and BB power spectra. (If information on the temperature is absent, one merely drops the \( X = T \) entries of the correlation matrix above.) If one believes that there is no foreground contamination, the cosmological signal on small scales should be dominated by E-mode polarization, so one can set \( Q_{B} = 0 \) in evaluating the likelihood.

9. DISCUSSION

The theoretical study of CMB anisotropies has advanced considerably in recent years, both in terms of theoretical predictions and comparison with observational data. In this paper we have developed some of the formalism necessary for the analysis of interferometer data such as will be returned by VSA, DASI, CBI, and VLA in the modern language of anisotropy (temperature and polarization) power spectra. Although the fundamental entities measured by an interferometer differ considerably from those measured in single-dish experiments, much of the analysis can be presented in an analogous manner, including power spectra, window functions, and the like. This allows one to use directly the sophisticated analysis techniques that already exist for single-dish experiments.

The formalism presented here is based on the "flat-sky" approximation and is thus best suited to the study of small-scale anisotropies. In this regime, interferometers can provide high-sensitivity measurements of the high-\( \ell \) peaks in the angular power spectrum, effects of gravitational lensing, the damping tail of the anisotropies and protogalaxy formation, and second-order anisotropies. For CBI the cosmological signal would be higher if the universe had significant spatial curvature, since this would shift the features in the angular power spectra to smaller angular scales; however, for both VSA and DASI significant signal is expected simply based on existing measurements.

Several outstanding problems remain, all associated with extending the framework of this paper to larger angular scales, where the curvature of the sky and the full three-dimensional nature of the FTs become important. We intend to return to these issues in a future paper.

APPENDIX A

FLAT-SKY APPROXIMATION

The derivation of the correlation function, equation (7) in the text, is sufficient for the purposes of this paper. However, sometimes it is convenient to treat the sky as flat and replace spherical harmonic sums with FTs at the temperature (rather than the two-point function) level. We give some details of this in this Appendix; the reader is also referred to Zaldarriaga & Seljak (1997).

In the flat sky approximation we replace the spherical polar coordinates \((\theta, \phi)\) by radial coordinates on a plane: \( r \equiv 2 \sin \theta / 2 \approx \theta \) and \( \phi \). We can exchange our indices \((\ell, m)\) for a two-dimensional vector \( \mathbf{I} \) with length \( \ell \) and azimuthal angle \( \phi_{r} \) and define

\[
a(l) = \sqrt{\frac{4\pi}{2\ell + 1}} \sum_{m} i^{-m} a_{lm} e^{im\phi_{r}} ,
\]

with \( a(-l) = a^{*}(l) \). We now expand our temperature field in terms of multipole moments as usual with

\[
\sum_{lm} a_{lm} Y_{lm} = \frac{1}{2\pi} \int d\ell \, d\phi_{r} \, a(l) \sqrt{\frac{2\ell + 1}{4\pi}} \sum_{m} i^{m} e^{-im\phi_{r}} Y_{lm} ,
\]

where we have replaced the sum over \( \ell \) with an integral.

Writing the sum over \( m \) in terms of \( m < 0 \) and replacing the associated Legendre polynomials with Bessel functions using Gradshteyn & Ryzhik (1980, eq. 8722[2]), we have

\[
\sum_{lm} a_{lm} Y_{lm} \approx \frac{1}{(2\pi)^{2}} \int d\ell \, d\phi_{r} \, a(l) \left[ J_{0}(\ell \theta) + 2 \sum_{m=1}^{\infty} i^{m} J_{m}(\ell \theta) \cos m(\phi_{r} - \phi_{l}) \right] .
\]
The term in square brackets is simply the Rayleigh expansion of a plane wave, so we finally obtain, writing \( l = 2\pi u \),

\[
\frac{\Delta T}{T}(x) = \sum_{rm} a_{rm} Y_{rm} \approx \int d^2 u \, a(u) \exp(2\pi i u \cdot x), \tag{A4}
\]

and of course the two-point function for \( a(u) \) is diagonal:

\[
\langle a^*(u) a(w) \rangle = C_r \delta^{(2)}(u - w). \tag{A5}
\]

These expressions can then be used to derive equation (8), although the derivation presented above equation (7) uses more controlled approximations. The generalization of equation (A4) to polarization is presented in § 8.

REFERENCES

Banday, A., & Wolfffendale, A. W. 1991, MNRAS, 248, 705
Becker, R. H., White, R. L., & Heliland, D. J. 1996, ApJ, 450, 559
Bennett, C. L., et al. 1992, ApJ, 396, L7
Bennett, C. L., Turner, M., & White, M. 1997, Phys. Today, Nov, 32
Bond, J. R. 1995, Phys. Rev. Lett., 74, 4369
Bond, J. R., Efstathiou, G., & Tegmark, M. 1997, MNRAS, 291, L31
Bond, J. R., & Jaffe, A. H. 1996, in Microwave Background Anisotropies, ed. F. Bouchet (Gif-sur-Yvette: Editions Frontières), 197
Bond, J. R., Jaffe, A. H., & Knox, L. 1998, Phys. Rev. D, 57, 2117
Brandt, W. N., Lawrence, C. R., Readehead, A. C. S., Pukhlanathan, J. N., & Fiola, T. M. 1994, ApJ, 424, 1
Bunn, E. F. 1995, Ph.D. thesis, Univ. California, Berkeley
Bunn, E. F., Fisher, K. B., Hoffman, Y., Lahav, O., Silk, J., & Zaroubi, S. 1994, ApJ, 432, L79
Bunn, E. F., Hoffman, Y., & Silk, J. 1996, ApJ, 464, 1
Bunn, E. F., Scott, D., & White, M. 1995, ApJ, 441, L9
Bunn, E. F., & Sugiyama, N. 1995, ApJ, 446, 49
Bunn, E. F., & White, M. 1996, ApJ, 480, 6
Carlstrom, J. E., Joy, M., & Grego, L. 1996, ApJ, 461, L59
Church, S. E. 1995, MNRAS, 272, 551
Cornwell, T. 1989, in ASP Conf. Ser. 6, Synthesis Imaging in Radio Astronomy, ed. R. Perley, F. Schwab, & A. Bridle (San Francisco: ASP), 277
Cooray, A. R., et al. 1997, A&AS, 191, 1906
de Oliveira-Costa, A., Kogut, A., Devlin, M. J., Netterfield, C. B., Page, L. A., & Wollack, E. J. 1997, ApJ, 482, L17
Dodelson, S. 1997, ApJ, 482, 577
Draine, B. T., & Lazarian, A. 1998, ApJ, 494, L19
Fiszen, D. J., Cheng, E. S., Gales, J. M., Mather, J. C., Shafer, R. A., & Wright, E. L. 1996, ApJ, 473, 576
Fomalont, E. B., et al. 1984, ApJ, 277, L23
Fomalont, E. B., Partridge, R. B., Lowenthal, J. D., & Windhorst, R. A. 1993, ApJ, 404, 8
Franceschini, A., et al. 1991, A&AS, 89, 285
Gorski, K. M., Banday, A. J., Bennett, C. L., Hinshaw, G., Kogut, A., Smoot, G. F., & Wollack, E. L. 1996, ApJ, 464, L11
Gradsteyn, I. S., & Ryzhik, I. M. 1980, Table of Integrals, Series, and Products (New York: Academic)
Halverson, et al. 1998, in ASP Conf. Proc. 141, Astrophysics from Antarctica, ed. G. Novak & R. Landsberg (San Francisco: ASP), in press
Heavens, A. F., & Petri, N. A. 1997, MNRAS, 285, 863
Hobson, M. P., Lasenby, A. N., & Jones, M. 1995, MNRAS, 275, 863
Hobson, M. P., & Mageuio, J. 1996, MNRAS, 283, 1133
Hogan, C. J., & Partridge, R. B. 1989, ApJ, 341, L29
Hu, W., Seljak, U., White, M., & Zaldarriaga, M. 1998, Phys. Rev. D, 57, 3290
Hu, W., & White, M. 1997a, NewA, 2, 323
—. 1997b, Phys. Rev. D, 56, 596
Jones, M. E. 1996, in Microwave Background Anisotropies, Proc. 26th Moriond Astrophysics Meeting, ed. F. R. Bouchet (Gif-sur-Yvette: Editions Frontières), 161
—. 1997, PPEUC Proc, Cambridge, April 7–11
Jungman, G., Kamionkowski, M., Kosowsky, A., & Spergel, D. N. 1996, Phys. Rev. D, 48, 3502
Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, Phys. Rev. Lett., 78, 2058
Savegnago, R., Ekers, R. D., Sinclair, M., & Silk, J. 1993, MNRAS, 263, 416
Subrahmanyan, R., Kesteven, M. J., Ekers, R. D., & Silk J. 1998, MNRAS, 300, 255
Subrahmanyan, R., Kesteven, M. J., Ekers, R. D., & Silk J. 1998, MNRAS, in press
Tegmark, M. 1997a, preprint (astro-ph/9712038)
—. 1997b, preprint (astro-ph/9710269)
Stompor, R., & Efstathiou, G. 1998, preprint (astro-ph/9805294)
Subrahmanyan, R., Ekers, R. D., Sinclair, M., & Silk, J. 1993, MNRAS, 263, 416
Subrahmanyan, R., Kesteven, M. J., Ekers, R. D., & Silk J. 1998, MNRAS, in press
White, M. 1998, Phys. Rev. D, 57, 5273
White, M., & Dunn, E. F. 1995a, ApJ, 443, L53
—. 1995b, ApJ, 450, 477 (erratum 460, 1071)
Wright, E. L. 1998, ApJ, 496, 1
Zaldarriaga, M. 1997, preprint (astro-ph/9702971)
Zaldarriaga, M., & Seljak, U. 1997, Phys. Rev. D, 55, 1830
Zaldarriaga, M., Spergel, D. N., & Seljak, U. 1997, ApJ, 488, 1
Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 457, 50