Senior Thesis Report: CMB Weak Lensing
Liam Fitzpatrick
Advisor: Bruce Winstein
Co-Reader: Kendrick Smith University of Chicago
May 24, 2004

Contents
1 Introduction 2
2 $C_\ell$ Spectra 2
3 Polarization 3
  3.1 Recombination ........................................ 3
  3.2 Q and U ............................................. 4
  3.3 E-modes and B-modes .................................. 4
  3.4 Covariance in Physical Space ......................... 5
4 Lensing 5
5 Simulations and Analysis 6
  5.1 Maps and Covariance Matrices ......................... 6
  5.2 Experimental Concerns: Beamsizes, Noise, Scan Patterns, and Binning .. 7
  5.3 Making Lensed Simulations ............................ 8
  5.4 Likelihoods ........................................... 9
  5.5 $l$ Ranges and Map Ranges ............................ 10
  5.6 Bands ............................................... 10
6 Results 11
  6.1 Variances ............................................ 11
  6.2 Correlations ......................................... 12
7 Acknowledgments 12
8 References 12

List of Figures
1 Q and U local definitions ................................. 5
2 E and B parity .......................................... 13
3 Q and U global definitions .............................. 14
4 $C_\ell$ Spectra .......................................... 15
1 Introduction

Cosmology has famously entered an era of unprecedented precision, and there is every reason to believe the trend of quickly improving measurements will continue. Observations have moved far beyond detecting the Cosmic Microwave Background (CMB) to making precise measurements of the details of its tiny inhomogeneities (called anisotropies) to within several arcminutes on the sky. Observations are now trying to do the same thing for the polarization of the CMB, which is roughly an order of magnitude smaller, and within the next few years may be able to detect gravity effects on that polarization. These “weak lensing” effects are predicted to be roughly another order of magnitude smaller, on the order of tenths of μK, and are simply a much weaker version of the well-known strong lensing that can cause multiple images. Since the effect is caused by the gravitational potential along the CMB photons’ paths, studying it can give us information about the distribution and nature of the matter in the universe. Currently, the QUIET group, a collaboration of several CMB groups from Chicago, Princeton, Miami, Harvard, Caltech, JPL, and others, is preparing for a five-year observation that hopes to detect this effect among other things.

In this paper, I will consider the effect of the weak lensing on the statistics of the CMB polarization. Sections 2-4 will describe the theory behind the CMB polarization and lensing. Sections 5.1-5.3 will describe a few practical issues related to actual observations with an eye toward past CAPMAP observations and future QUIET ones, and how to simulate data from such observations. Sections 5.4-5.6 will describe how I analyze real or simulated data. Finally, in section 6, I will present my results on the statistics of the lensed CMB polarization for different experimental conditions.

2 $C_l$ Spectra

Since most of the language in this report will revolve around “$C_l$ spectra”, I will begin by defining them and trying to give some intuition on what they mean. Simply put, the $C_l$ spectrum of a field is its expected variation on scales proportional to $l^{-1}$. More precisely, we can decompose any scalar field $X(\hat{n})$ on the sky into

$$X(\hat{n}) = \sum_{lm} a_{lm} Y_{lm}(\hat{n})$$

(1)

or, more generally, any spin-$s$ field $sX(\hat{n})$ into
\[ sX(\hat{n}) = \sum_{lm} a_{lm} sY_{lm}(\hat{n}) \] (2)

where the \( Y_{lm} \)'s and \( sY_{lm} \)'s are spherical harmonics and spin-weighted spherical harmonics, respectively\(^1\). The \( l \)'s are called \"multipoles\", and looking at a field in \"multipole-space\" is very much like looking at a function on the \( xy \)-plane in Fourier-space. For a given \( l \), \( Y_{lm} \) varies on order of \( \frac{\pi}{l} \) radians or about \( \frac{100^\circ}{l} \), that is, that is the width of its \"wiggles\". When we are talking about a random scalar field, so that \( X \) is drawn from some distribution of possible fields, then the \( a_{lm} \)'s must also be random, and we can ask what their statistics are. No matter what their distributions look like, we can always define \( C_{lm} = \langle a_{lm}a_{lm}^* \rangle \).\(^2\) If these statistics do not have a preferred direction, as is true of the CMB, then \( C_{lm} = C_{lmm'} \) for all \( m, m' \), and we drop the \( m \) subscript.

In fact, in most of the cases important here, we can simplify even further. Usually, we will also have

\[
\begin{align*}
\langle a_{lm} \rangle & = 0 \\
\langle a_{lm}a_{l'm'}^* \rangle & = \delta_{ll'}\delta_{mm'}C_l
\end{align*}
\] (3) (4)

Consider for a moment what this means for an experiment trying to measure \( X \). Whereas in a typical experiment, for example one measuring the temperature of neutron stars, we are interested mostly in the average or expected value, and the spread in those values is less important and may even be undesirable. Here, however, the expected value is totally uninteresting - in physical and in multipole-space, it is zero everywhere! What \emph{is} interesting is the variation in the value. Moreover, since \( a_{lm} \)'s with the same \( m \) subscript are uncorrelated and have the same statistics, they are essentially drawn from the same distribution. Thus, a perfect measurement of a single \( X \) yields a perfect measurement of \( 2l + 1 \) \( a_{lm} \)'s for each \( C_l \), and their sample variance ought to be close to \( C_l \), the variance of the distribution from which they are drawn. This is very fortunate, since in practice there is only one CMB, and thus there is only one \( X \) to measure. If \( X \)'s \( C_l \) spectrum is large at \( l' \), then it is more likely that \( a_{l'm} \) will be large for \( m = -l', \ldots, l' \), and thus \( X \) will be likely to have large fluctuations on scales around \( \frac{100^\circ}{l'} \).

3 Polarization

3.1 Recombination

The CMB was initially polarized by Thompson scattering of photons off of free electrons in the early universe. When the universe was hot and dense, there were far more free electrons, and this scattering occurred frequently. Once the universe cooled enough for electrons and protons to form stable hydrogen at an age called \"recombination\", this scattering became far more rare, and most CMB photons have, to first-order, maintained the same polarization since

\(^1\)Fields with spin-weight \( s \) gain a phase factor \( e^{is\psi} \) when the basis axes rotate by an angle \( \psi \), and thus a rotation of \( 2\pi/s \) has no effect.

\(^2\)In this paper, \( \langle \ldots \rangle \) denotes ensemble average, so that \( \langle xy^* \rangle = Cov(x,y) \).
that time. For this reason, that time is also called the “surface of last scattering”. Because the
radiation in the early universe is not completely isotropic but instead has small variations, the
emitted photons were more polarized in some directions than in others; since these variations
are quadrupolar, the polarization is linear [1]

3.2 Q and U

Since polarization is the same if it is rotated 180°, it is a spin-2 quantity and needs spin-2
basis vectors. We call these basis vectors Q and U, and the convention for defining them globally
is shown in Figure 3. It is important, however, to realize that we are free to define them rotated
differently at every point. We write, as in equation 2,

\[(Q + iU)(\hat{n}) = \sum_{lm} a_{2,lm} 2 Y_{lm}(\hat{n}), \quad (Q - iU)(\hat{n}) = \sum_{lm} a_{-2,lm} -2 Y_{lm}(\hat{n})\]  

(5)

3.3 E-modes and B-modes

It is useful to decompose the polarization into the spin-2 equivalent of its curl-free and
divergence-free components; these are called E-modes and B-modes respectively, in analogy with
electromagnetism. As shown for the simple case \(l = 2, m = 0\) in figure 2, E-modes act the same
under parity as a scalar field, whereas B-modes gain an extra sign under parity transformations.
Since Thompson scattering results from the temperature field, it can only create polarization
with the former parity and thus only creates E-modes but not B-modes. Consequently, B-modes
ought to represent only lensing and primordial gravity waves.

In multipole space, we simply take \(a_{lm}^E = -(a_{2,lm} + a_{-2,lm})/2\) and \(a_{lm}^B = -(a_{2,lm} - a_{-2,lm})/2i\)
[2], as well as \(X_{1,lm}(\hat{n}) = (2 Y_{lm} + 2 Y_{lm})/2\) and \(X_{2,lm}(\hat{n}) = (2 Y_{lm} - 2 Y_{lm})/2\). This gives

\[Q(\hat{n}) = - \sum_{lm} a_{E,lm} X_{1,lm}(\hat{n}) + i a_{B,lm} X_{2,lm}(\hat{n}) \quad \quad U(\hat{n}) = - \sum_{lm} a_{B,lm} X_{1,lm}(\hat{n}) - i a_{E,lm} X_{2,lm}(\hat{n})\]  

(6)

In practice, we do not measure polarization on the entire sky, and thus the decomposition
into E- and B-modes is not unique and depends on ambiguous boundary conditions at the edge
of the area we observe. This leads to what is called “E-B leakage”, meaning that an experiment
measuring a pure E pattern on an incomplete patch of the sky will see some B pattern as
well, and vice versa. Since gravitational effects and therefore B-modes are much smaller than
Thompson scattering effects and E-modes, an experiment on a small patch runs the risk of most
of its B-mode pattern coming from “leaked” E-modes. This effect increases with the ratio of
boundary length over survey area [3].

Figure 4 shows log-plots of \(C_l^E = \langle a_{lm}^E a_{lm}^E \rangle\) and \(C_l^B = \langle a_{lm}^B a_{lm}^B \rangle^3\), as well as the \(C_l\) spectrum
for temperature, which I will not consider.

\(^3\)These spectra do not include lensing.
3.4 Covariance in Physical Space

The best we can do theoretically is calculate the statistics of \( Q \) and \( U \). Since, as mentioned, their expected values are zero everywhere, the next thing we are interested in is their covariances. A cosmological theory will predict \( C^E_l \) and \( C^B_l \). For any two points on the sky \( \hat{n}_1 \) and \( \hat{n}_1' \), we could take the above formulae for \( Q(\hat{n}) \) and \( U(\hat{n}') \) and plug them into \( \langle Q(\hat{n})U(\hat{n}') \rangle \), getting a complicated formula with sums of products of \( X_{i,lm} \)'s, but in practice the formulae simplify considerably if we redefine \( Q \) and \( U \) in a local reference frame depending on the two points between which we are calculating the covariance. This local frame is shown in Figure 1, where the two big dots represent \( \hat{n} \) and \( \hat{n}' \), and the direction of \( +Q \) is along the line between them. First, we do the calculation in this frame, and then we rotate to the global reference frame in Figure 3.

![Figure 1: Q and U local definitions](image)

In the local reference frame, the formulae are [4]

\[
\langle Q(\hat{n})Q(\hat{n}') \rangle = \sum_l \left( \frac{2l+1}{4\pi} \right) [F^E_l(z)C^E_l - F^B_l(z)C^B_l] \\
\langle Q(\hat{n})U(\hat{n}') \rangle = 0 \\
\langle U(\hat{n})U(\hat{n}') \rangle = \sum_l \left( \frac{2l+1}{4\pi} \right) [F^B_l(z)C^B_l - F^E_l(z)C^E_l]
\]

where \( z = \hat{n} \cdot \hat{n}' \) is the cosine of the angle between the two points and the \( F \) functions can be written in terms of associated Legendre polynomials. If we want only E-mode or B-mode patterns, we leave out \( C^B_l \) or \( C^E_l \), respectively. To get to the global frame, we need only to find the necessary rotation angle \( \psi \) between frames and then transform according to \( Q' = Q \cos(2\psi) + U \sin(2\psi), \ U' = U \cos(2\psi) - Q \cos(2\psi) \) [4].

4 Lensing

As the CMB photons travel from the surface of last scattering, they pass through a gravitational potential due to the matter distribution in the universe, thus bending their paths slightly. From earth, we see this three dimensional potential effectively collapsed onto the two-dimensional sky, and we call this two-dimensional potential \( \Phi \). It is a random scalar field, and so we can use the
formalism of the $C_l$ spectra to describe it. Integrating over the photons' paths back to $z_r \approx 1100$, the redshift at last scattering, one has

$$C_l^{\Phi} = \frac{(2\pi)^2}{l^3} \int_{z_0=0}^{z_r=1100} \frac{c dz'}{H(z')} D \left( -2 \frac{D(z_r) - D(z')}{D(z_r)D(z')} \right)^2 \Delta_\Phi^2(k = l/D, z')$$

(10)

where $c$ is the speed of light, $H(z) = H_0 \sqrt{\Omega_m(1 + z)^3 + \Omega_\Lambda}$ is the Hubble constant, $\Delta_\Phi^2(k, z)$ is the variance at redshift $z$ of the lensing at wavenumber $k$, $\Omega_m$ and $\Omega_\Lambda$ are the fractional mass densities of matter and dark energy, and

$$D = \int_{z_0}^{z} \frac{c dz'}{H(z')}$$

(11)

is the comoving distance. The polarization then changes simply as

$$Q'(\hat{n}) = Q(\hat{n} + \nabla \Phi(\hat{n}))$$

(12)

$$U'(\hat{n}) = U(\hat{n} + \nabla \Phi(\hat{n}))$$

(13)

Since it is in fact $\nabla \Phi$ that appears here, we consider its statistics rather than those of $\Phi$. They are, in a local reference frame, [5]

$$\langle \nabla_T \Phi(\hat{n}) \nabla_T \Phi(\hat{n}') \rangle = \sum_l \frac{2l + 1}{4\pi} C_l^{\Phi} \left( -z P_l'(z) + l(l + 1) P_l(z) \right)$$

(14)

$$\langle \nabla_T \Phi(\hat{n}) \nabla_N \Phi(\hat{n}') \rangle = 0$$

(15)

$$\langle \nabla_N \Phi(\hat{n}) \nabla_N \Phi(\hat{n}') \rangle = \sum_l \frac{2l + 1}{4\pi} C_l^{\Phi} P_l'(z)$$

(16)

where $T$ is the direction from $\hat{n}$ to $\hat{n}'$, $N$ is the directions perpendicular to $T$, $z = \hat{n} \cdot \hat{n}'$ as before, and $P_l$ and $P'_l$ are Legendre Polynomials and their derivatives, respectively. In Figure 1, $T$ points along $+Q$ from $\hat{n}$ to $\hat{n}'$, while $N$ points along $-Q$, chosen to be up or down so that it is 90° clock-wise to $T$.

One important consequence of equations (12) and (13) is that the $a_{lm}^E$'s and $a_{lm}^B$'s are no longer uncorrelated at different $l$ and for E-B. Later, I will look more closely at this change in statistics with simulations.

5 Simulations and Analysis

5.1 Maps and Covariance Matrices

After observing the sky for hundreds of hours, the CAPMAP group turns its data into a map with an average measurement in each of many bins on the sky. Two possible binning patterns are shown in Figure 5, the first of which I call the "rectangular" binning pattern and the second of which I call the "cap" binning pattern. In all my simulations, I use the rectangular binning with 20 bins on a side, and thus 400 total bins. A map of binned data is just a vector

6
where \( x_i \) can be a measurement of \( Q, U \), or some linear combination of them in bin \( i \). For simplicity, I will assume that we have a measurement of \( Q \) and \( U \) for every bin.

The map then has a theory covariance matrix associated with it, defined by

\[
\mathbf{F} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\]

\[(17)\]

where we know the right-hand side from equations (7)-(9). We can also calculate \( C_{T,i} \) for a given \( l' \) by taking only the terms in the sums in (7)-(9) with \( C_{l'} \), and likewise for \( C_{T,l}^B \).

5.2 Experimental Concerns: Beamsizes, Noise, Scan Patterns, and Binning

There are a few parameters to consider about the maps we make in practice. The first of these is the beamsize. In reality, we cannot sample the polarization at a single point on the sky; at best, we will still have contributions to our signal from nearby parts of the sky within the resolution of our telescope. The weight of those contributions will be a 2-dimensional gaussian centered on the point we are trying to measure, and the full-width half-maximum (FWHM) of this gaussian is our “beamsize” \( \theta_{FWHM} \) and is effectively the smallest scales that our telescope is capable of viewing. This damps the \( C_l \) spectrum we detect at high \( l \) (remember that \( l \) corresponds to fluctuations on scales of \( \theta \sim l^{-1} \) ) by a factor of \( B_l = \exp[-l(l+1)\sigma^2] \), where \( \sigma = \theta_{FWHM}/[8 \ln(2)]^{1/2} \) [2]. The beamsize for our experiments is \( \theta_{FWHM} = 0.063^\circ \). Figure 4 shows the effect of our beamsize on the \( C_l \)'s we can measure - notice that its effect only become important around \( l = 1000 \).

The second experimental consideration is instrument noise. The noise \( n \) (that is, the standard error) for any measurement at a point on the sky will be uncorrelated with the noise at any other point on the sky. The noise also goes down as we spend more time sampling the polarization at a particular point according to \( n \propto 1/\sqrt{t} \). Thus, since we only have a finite amount of time to take data, we have to compromise between covering a large part of the sky and having larger errors on each point versus covering a smaller part of the sky and having smaller errors on each point. Later, I shall consider the important question of what sky coverage is the best compromise between these two effects. Since the noise in different bins is uncorrelated, the noise covariance matrix is

\[
N_{ij} \equiv \langle n_in_j \rangle = n_i^2 \delta_{ij}
\]

\[(19)\]

The amount of time that we spend in each bin also depends on the motion of our telescope, which I refer to as our “scan pattern”. The first scan pattern I will consider, the “azimuth scan” or “az-scan”, is where we point our telescope at the sky and move it back and forth in azimuth. The second, the “ring scan”, is where we move our telescope in a small ring on the sky. As the earth turns, both of these patterns sweep out a larger area, as shown in Figure 6.
This affects the analysis by changing the noise in each pixel according to the amount of time each scan spends in it. To calculate this time weight, I wrote a monte carlo program that chooses a random time of day equally weighted throughout the day and a random point on the scan (on the line in the az-scan or the ring in the ring-scan) equally weighted on the scan. It takes $10^7$ such random points, and uses the number of random points that fall in a bin as the time weight for that bin. The variance, or noise-squared, of that bin is just inversely proportional to that weight. The time weight for the az-scan is shown in Figure 7 and for the ring-scan in Figure 8. I also use a uniform weighting system, which has the same noise in every pixel. In every case, I normalize the noise so that all three scans have the same amount of total time, since in real life changing our scan unfortunately does not gain us extra observing time. The normalization is such that, if we were to spend all of our time in a single bin, the noise in that bin would be 0.1 $\mu K$; we refer to this by saying the map has $senvac = 0.1\mu K$. In the uniform-noise scan, where we divide time equally over all 400 bins, the noise in each bin is therefore $senvac \times (\sqrt{\frac{t_{\text{bin}}}{t_{\text{total}}}})^{-1} = 0.1\mu K \times \sqrt{400} = 2.0\mu K$.

The scan strategy will also affect the results simply due to the geometry of the scan. The most obvious effect is the possible introduction of extra boundary conditions, such as the inner hole in each scans in Figure 6. Since we want map covariance that minimizes the boundary-to-area ration, these holes are clearly bad. More generally, any non-uniformity in the noise will have a boundary-condition-like effect. In a sense, a boundary is essentially nothing more than an area with infinite noise. Thus, we would expect the best measurement to come from the “uniform noise” scan, where the noise in each bin is exactly the same. In practice, it is not so easy to get uniform noise everywhere, since we a) want our telescope to move back and forth at a near-constant rate, and b) we want to come through bins at different angles in order to avoid degeneracy problems. I shall not go into either of these interesting problems here.

How we choose to bin the data can have an effect on our final answer if we bin poorly. If we bin so coarsely that we average too much data together, we will throw out important information (clearly, if we made our whole map just one big bin, we would have very small noise per bin on a useless map). However, the more bins we make, the larger our data vector is and the longer the data analysis takes; since the analysis is an $O(N^3)$ process, this is not an idle concern. In general, we do not start to throw out information until the bins become larger than our beamsize, as we would expect.

5.3 Making Lensed Simulations

Once we have the total covariance matrix $C = C_T + N$, it is straightforward to make simulations of the lensing potential and of unlensed polarization maps. To make simulations of the lensed polarization maps, one could write a program to implement equations (12) and (13), interpolating to get the values for the polarization between the grid spaces. However, as suggested by Kendrick my co-reader, it is easier to shift the points $\hat{n} \rightarrow \hat{n} + \nabla \Phi$ of the map, calculate a new covariance matrix, and simulate the polarization from that. This is not as straightforward as it looks at first sight. What exactly does it mean to add $\hat{n}$ and $\nabla \Phi$? The former is a point on the sphere, and the latter is a vector in 2-d space - for the mathematically-minded, $\hat{n} \in S^2$, and

---

4That is, one that relies on simulations.
\( \nabla \Phi \) is in the tangent space \( T_hS^2 \). We add them as follows.

Let \( \hat{n} = (\theta, \phi), \nabla \Phi = (\nabla_x \Phi, \nabla_y \Phi) \) and assume for simplicity (and by rotational symmetry without loss of generality) that \( \phi = 0 \). If \( \hat{x}, \hat{y}, \) and \( \hat{z} \) are the cartesian unit vectors, then this means \( \hat{n} \) points in the \( xx-plane \) and perpendicular to \( \hat{y} \). We should interpret \( (\theta, 0) + (\nabla_x \Phi, \nabla_y \Phi) \) as the point moved in the direction in which \( (\nabla_x \Phi, \nabla_y \Phi) \) points for \( |(\nabla_x \Phi, \nabla_y \Phi)| = \sqrt{\nabla_x \Phi^2 + \nabla_y \Phi^2} \) radians along a great circle. To accomplish this calculation, I take the great circle passing through the NCP and \( \hat{n} \), rotate that by \( \alpha = \arctan(\nabla_y \Phi / \nabla_x \Phi) \) to get the direction in which \( (\theta, \phi) \) points, and then rotate \( \hat{n} \) by \( |(\nabla_x \Phi, \nabla_y \Phi)| \) radians along that great circle. More precisely, as mentioned, \( \hat{y} \) is the normal vector to the great circle passing through the NCP and \( \hat{n} \). \( \hat{n} \) and \( \hat{y} \) are perpendicular, so together with \( \hat{y}' = \frac{\hat{n} \times \hat{y}}{|\hat{n} \times \hat{y}|} \), they form an orthonormal basis. Thus, rotating this great circle by \( \alpha \) means rotating

\[
\hat{y} \rightarrow \hat{y} \cos(\alpha) + \hat{y}' \sin \alpha \equiv \hat{y}_\alpha \tag{20}
\]

Similarly, rotating \( \hat{n} \) along this newly formed great circle means rotating

\[
\hat{n} + \nabla \Phi = \hat{n} \cos(|(\nabla_x \Phi, \nabla_y \Phi)|) + \frac{\hat{y}_\alpha \times \hat{n}}{|\hat{y}_\alpha \times \hat{n}|} \sin(|(\nabla_x \Phi, \nabla_y \Phi)|) \tag{21}
\]

5.4 Likelihoods

The likelihood distribution is the probability of the data set given a certain theory (or, equivalently, given a certain covariance matrix), and we want to know for what theory this is at a maximum. In practice, it is more convenient to consider the log-likelihood; maximizing the log of the likelihood is equivalent to maximizing the likelihood itself. The likelihood is given by

\[
\mathcal{L}(\alpha) \sim [\det C]^{-1/2} \exp\left[-\frac{1}{2} \mathbf{x}^T C^{-1} \mathbf{x}\right] \tag{22}
\]

where \( C = \alpha C_T + N \), the sum of the theory covariance matrix and the noise covariance matrix. The \( \alpha \)-dependence of \( \mathcal{L} \) is in \( C = C(\alpha) \). We can transform into the eigenbasis for \( \mathbf{C} \), so that \( \mathbf{C} \rightarrow \text{diag}(\gamma_i), \mathbf{x} \rightarrow \tilde{\mathbf{x}} \), where \( \gamma_i \) are the eigenvalues of \( \mathbf{C} \). In this basis, we can calculate the log-likelihood by the brute force method [6]:

\[
\log [\mathcal{L}(\alpha)] = -\frac{1}{2} \sum_{i=1}^{n} \left[ \log(\gamma_i) + \frac{x_i^2}{\gamma_i} \right] \tag{23}
\]

where the \( \alpha \)-dependence is in \( \gamma_i \), the eigenvalues of \( C(\alpha) \). I then search for the parameters \( \alpha \) that maximize \( \log \mathcal{L} \). Unfortunately, this is a very lengthy procedure for large data sets, since it requires transforming into a new eigenmode basis for each evaluation of the log-likelihood. In practice, I use the "signal-to-noise eigenmode" method. This amounts to transforming to a new basis where \( N = I \) and \( C_T \) is diagonal. Then

\[
C = \text{diag}(1 + \alpha \gamma_i) \tag{24}
\]

and
\[ \log [L(\alpha)] = -\frac{1}{2} \sum_{i=1}^{n} \left[ \log(1 + \alpha \gamma_i) + \frac{x_i^2}{1 + \alpha \gamma_i} \right] \] (25)

where \( \gamma_i \) and \( z_i \) are unchanged for different values of \( \alpha \). I should emphasize here that this change of basis is only possible if the theory covariance matrix is positive-definite \(^5\), which all true covariance matrices are by a simple theorem of statistics. However, if \( \alpha \) falls far enough below zero, then \( C = \alpha C_T + N \) will inevitably no longer be positive-definite - thus, I constrain \( \alpha \) to be non-negative. The reasons that the likelihood would peak at negative signal would be that one overestimated the noise matrix or that the map happened to have extremely low signal. This will be important later.

### 5.5 l Ranges and Map Ranges

Since \( l \) corresponds to fluctuations of a certain size in physical space, the size of our map limits the range that we can probe effectively. Just as a grid in physical space with length \( L \) and grid spacing \( \Delta x \) will only have well-defined fourier transforms from \( k = L^{-1} \) to \( k = (\Delta x)^{-1} \), similarly our map will only probe at best a range of \( l \). Since my simulated maps here have 20 pixels on a side, the grid spacing is \( \Delta x = \frac{L}{20} \) and they will probe a range of

\[ 100 \frac{1^\circ}{L} \lesssim l \lesssim 200 \frac{1^\circ}{L} \] (26)

Table 1 gives the map sizes I use and their corresponding \( l \) range; the total range is about \( l = 13 \) to \( l = 2000 \).

### 5.6 Bands

In order to recover \( E \) and \( B \) at different multipoles, one can introduce extra free parameters for each multipole \( l \):

\[ C = N + \sum_l \alpha_l^E C_{T,l}^E + \sum_l \alpha_l^B C_{T,l}^B \] (27)

However, since we want to probe up to around \( l = 2000 \), that would give a prohibitive number of parameters to minimize over. In practice, we bin the multipoles into bands:

\[ C_{b_i}^{E,B} = \sum_{l \in b_i} C_{T,l}^{E,B} \] (28)

\[ C' = N + \sum_i \alpha_i^E C_{b_i}^E + \sum_i \alpha_i^B C_{b_i}^B \] (29)

Ideally, given lots of computer time, one would choose several bands. In order to get more simulations, I took only two bands each for \( E \) and \( B \), spaced to have equal width in logarithm-space, from \( l = 40 \) to \( l = 2000 \), so that band 1 goes from \( l = 40 - 282 \) and band 2 goes from

\(^5\)A matrix is said to be positive-definite if all of its eigenvalues are positive.
\( l = 283 - 2000 \). Notice that the smallest map size, 1°, will see only about half of the first band in logarithm space \(^6\). Table 1 tabulates more of these coverage values.

Since I fit a multiplier to \( C_{b_i}^{E,B} \), I am essentially assuming that I know the shape of the \( C_l \) spectra but not their amplitudes. If the shape I use happens to be correct, then all is well. However, the worse the shape differs from the actual \( C_l \) spectra, the more adverse effects will come in, in particular E-B leakage. Since I am recovering E and B for lensed maps, I use \( C_l^E \) and \( C_l^B \) spectra that include lensing. While equation (4) is no longer true with lensing, that does not prevent us from defining \( C_l \) spectra by \( C_{l \text{ lensed}} = \langle a_{lm}^{\text{ lensed}} a_{lm}^{* \text{ lensed}} \rangle \) as always. The program I use\(^7\) to get my \( C_{l \text{ lensed}}^{E,B} \) spectra in the first place also calculates these lensed spectra.

\section{Results}

\subsection{Variances}

We can now ask what map size gives the best errors. Looking at Table 1, we see that we have good coverage for both bands at around 2.0°. A simple way to estimate the best map size more accurately is from the fit values \( \alpha_i \) from all the simulations. For each scan and each map size, I made and analyzed about 25 simulations. Each of these analyses results in a value for each \( \alpha_i \), and so I can calculate the variance of the fit values for each one. I show these in Figure 9. First of all, one can see that the ring scan is almost as good as uniform noise, and both are better than the az-scan in the two lower \( l \) bands - however, in the two upper \( l \) bands, the variances are all nearly the same. Aside from the several statistical fluctuations, the trend seems to be that the variance is minimized around 2.0°-3.0°, which is close to what we would guess from our (very rough) estimate based on \( l \)-space coverage. From Table 1, we would also expect to see one other effect: the variance in the higher \( l \) band should start to rise rapidly as we approach 8.0°. This does indeed happen in both the E and B higher band, after mildly-sloped curves near map sizes of 2°. There is also the mysterious feature in some of the lower \( l \) bands that the variance decreases at the smallest maps. This represents the fact that I mentioned earlier that I constrain the fit parameter to be non-negative, so that it will be exactly zero if we see very little signal. Of the 25 fit values for B\(|l = 40 - 282|\) for each scan at a map size of 1.0°, only 5 were non-zero for uniform noise, only 2 were non-zero for the ring-scan, and none non-zero for the az-scan. In contrast, up only 0.5° more in the same band, 12 were non-zero for uniform noise, 3 were non-zero for the ring-scan, and 9 were non-zero for the az-scan.

Perhaps most important is the level of the variance in the higher B band. At its minimum, it reaches slightly below 1.0. 1.0 is the value that we ought to get as our fit value for \( \alpha \), since it is just a multiplier by the theory covariance matrix. In fact, we typically get about twice that, which should not be startling since as I mentioned \( C_{l \text{ lensed}} \) only includes the variance of the \( a_{lm}^{\text{ lensed}} \)'s but not higher moments. I show the actual average fit values for this higher band as *'s. Aside from one outlier, all three scan strategies should give about a 1-1.5σ detection of the higher B-band if none of these were pure B-modes. Unfortunately, it seems that about 75% of this detection comes not from lensing but from leaked E-modes. To see this, I tried running

\(^6(\log(282)-\log(100))/(\log(282)-\log(40)) = 0.53\)

\(^7\)CMBFAST, by M. Zaldarriaga and U. Seljak
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
<th># of Sims</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lensed</td>
<td>2.07</td>
<td>0.9266</td>
<td>25</td>
</tr>
<tr>
<td>Unlensed</td>
<td>1.5057</td>
<td>1.2074</td>
<td>9</td>
</tr>
</tbody>
</table>

the uniform-noise scan on a 3.0° map, but used $\nabla \Phi = 0.0$ everywhere for my lensing simulation. Table 6.1 compares the results with and without lensing, where because of time constraints the number of simulations is clearly far fewer than we would like. Of course, because of the small number of simulations, the amount of B-detection might be much more than $\frac{1}{2 \times 2.5} \approx 75\%$. A simple $t$-distribution test gives 90.1% confidence that the B-detection is not all E-leakage.

6.2 Correlations

We can also estimate the correlations between different bands by looking at the sample correlations from the simulations. I show these for uniform noise in a lego-style plot in Figure 10 and in a one-dimensional plot in Figure 11. The bottom line seems to be that the statistical fluctuations are too great to make any statement about the correlations themselves. Twenty-five simulations per mapslice simply is not enough to see small correlations, even if they were around the level of 0.4. While the process is the same to get 1000 simulations as it is to get 25, I unfortunately ran out of time before I could run more.

7 Acknowledgments

I am grateful to Takemi Okamoto and Wayne Hu for useful discussions and guidance. I would also like to thank my co-reader Kendrick Smith for his assistance with calculations, defining the problem, and suggestions on how to tackle it, as well as my advisor Bruce Winstein for teaching me about how to do good science and for at times having more confidence in me than I did myself. I am also grateful to generous support from the Grainger family and the Goldwater Foundation.

8 References

Figure 2: E and B parity
Figure 3: Q and U global definitions
Figure 4: $C_l$ Spectra, unlensed. The black line shows the ideal spectra, while the blue line shows the spectra including the effect of our beamsize. Damping from the beam is not significant up until $\log l \approx 3$. The asterisks in the right-most panels are the average of the best-fit values for the 25 simulations, which are essentially the signal detected.
Figure 5: Binning Examples

<table>
<thead>
<tr>
<th>Map width $L$ (°)</th>
<th>Lowest $l$</th>
<th>Highest $l$</th>
<th>% of band 1 covered</th>
<th>% of band 2 covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>100</td>
<td>2000</td>
<td>53</td>
<td>100</td>
</tr>
<tr>
<td>1.5</td>
<td>67</td>
<td>1333</td>
<td>73</td>
<td>79</td>
</tr>
<tr>
<td>2.0</td>
<td>50</td>
<td>1000</td>
<td>89</td>
<td>65</td>
</tr>
<tr>
<td>3.0</td>
<td>34</td>
<td>667</td>
<td>100</td>
<td>44</td>
</tr>
<tr>
<td>4.0</td>
<td>25</td>
<td>500</td>
<td>100</td>
<td>29</td>
</tr>
<tr>
<td>5.0</td>
<td>20</td>
<td>400</td>
<td>100</td>
<td>18</td>
</tr>
<tr>
<td>6.0</td>
<td>17</td>
<td>334</td>
<td>100</td>
<td>8.5</td>
</tr>
<tr>
<td>7.0</td>
<td>15</td>
<td>285</td>
<td>100</td>
<td>0.36</td>
</tr>
<tr>
<td>8.0</td>
<td>13</td>
<td>250</td>
<td>94</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 1: $l$ ranges and Map size
Figure 6: Ring and Az patterns
Figure 7: Az-scan time Weighting. All four images are of the same noise, but from different angles.
Figure 8: Ring-scan time Weighting. All four images are of the same noise, but from different angles.
Figure 9: Variances for all Bands. The x-axis is the length of the side of the map area in degrees, and the y-axis is the sample variance in $\mu{K}$ of the fit values for $\approx 25$ simulations each. In the last row, black is for uniform noise, blue is for the ring-scan, and red is for the az-scan.
Figure 10: Band Correlations for Uniform Noise. The number below each plot is the size of the map (e.g. 6.0°). The height of each bar is the correlation between the bands corresponding to its place on the x-axis and its place on the y-axis; band 1 is E(2-282), band 2 is E(283-2000), band 3 is B(2-282), band 4 is B(283-2000). Thus, all bars on the diagonal have height=1.0.
Figure 11: Band Correlations, Uniform Noise, 1-dim