Polychromatic map-making with asymmetric antenna patterns

Connor Mooney
Polychromatic Map-Making with Asymmetric Antenna Patterns

Connor Mooney
March 2019

Honors Thesis
Submitted to:

Physics Department
University of Richmond
Richmond, VA

Advisor: Dr. Emory Bunn
Abstract

Imaging telescopes with asymmetric antenna patterns that vary with wavelength can create time-ordered data that may be processed into multiple images corresponding to different bands of wavelengths from just a single set of scans. The imaging telescope named QUBIC has this property and is the inspiration behind this project. Our goal is to quantify, both statistically and analytically, the ability of these telescopes to perform such a reconstruction given different cases. In the case that the telescope is observing the full sky, we reconstruct our maps via a spherical harmonic basis. In this way, the reconstructed images are described as a set of spherical harmonic coefficients, whose properties can be analyzed and computed relatively easily. In the case that the telescope is only observing part of the sky, we must reconstruct maps with a value assigned to each discretized point in the sky, and thus more computation and analysis is required. In each case, we find eigenvectors in wavelength space that maximize the reconstructed signal-to-noise ratio, and use these to quantify the number of maps that can be reconstructed accurately.

Introduction

If we remove all the noise from our sun and our Galaxy, virtually all of the radiation left in the sky is from the cosmic microwave background (CMB). This ancient light dates back to the point of recombination when the universe was only 300,000 years old. Although much more energetic at the time, those photons have reached present day Earth with a wavelength of about 2mm. Being the oldest light observable, the CMB is capable of answering many fundamental questions of our universe. Therefore, it is important that our telescopes observing the sky are able to reproduce the most accurate and informative maps of the sky as possible.

My research is focused on analyzing the plausibility of reconstructing maps of the CMB over multiple wavelengths given our unique antenna pattern. An antenna pattern (or beam pattern) is a property of an imaging telescope, which characterizes how an incoming signal from the sky is converted to usable information. An antenna pattern gives a certain amount of weight to the signal it receives relative the the angle of separation away from the center of where the telescope is pointing. So, a telescope will scan across the sky, and at each unit of time, the antenna pattern is multiplied with the intensity of the photons of the sky and then averaged to get a single point of time-ordered data for a given wavelength (Quinn, 2018). After a large enough amount of time, a sufficient quantity of time-ordered data will be produced, and we should be able to reproduce a good reconstruction of the original sky maps with low noise.

Traditional telescopes will have a symmetric antenna pattern; however, our telescope (QUBIC), with its asymmetric pattern that varies with wavelength,
has the advantage of measuring the signal from the sky at multiple wavelength bands with just a single set of scans. In this case, asymmetric means that the beam pattern will vary with the change of the azimuthal angle, $\phi$. QUBIC stands for Q & U Bolometric Interferometer for Cosmology, which measures B-mode polarization of the CMB (Tartari, 2016). An interferometer is a telescope that employs the interference of incoming light to make measurements. A bolometric device measures the total intensity of light from its source. QUBIC’s antenna pattern is specifically composed of multiple Gaussian peaks under a wider virtual Gaussian envelope. The contour plot of the antenna pattern is shown below, where the center of each red circle is a local maximum of a Gaussian peak. Note that the central peak is the absolute maximum and represents the center of the antenna pattern. This pattern is projected onto the surface of the sphere that makes up the CMB.

![Figure 1: Top-down contour plot of QUBIC’s antenna pattern. (Stopovskiy, 2016)](image)
2 Monochromatic Case

First we begin by encapsulating our antenna pattern within a matrix that we call a pointing matrix, \( A \), in which each element of a row corresponds to the weight of the antenna pattern at that specific pixel of the sky, where each individual row corresponds to one orientation of the telescope at a point in time. We then steer the telescope over the sky to generate a series of time-ordered data.

\[
d_t = \sum_p A(R_t \vec{r}_p) s_p + n_t
\]

\( \vec{r}_p \) is the location of the \( p^{th} \) pixel, \( R_t \) is the rotation matrix indicating the position of the telescope at time \( t \). \( s_p \) is the signal from the sky at the \( p^{th} \) pixel, and \( n_t \) is the noise in the time domain.

It is important to note that when we are reconstructing a map in pixel-space, we use a simplified model for our antenna pattern. Smooth Gaussian peaks are replaced with a delta spikes placed at the position where each local maximum is located instead. The purpose of a smooth function for an antenna pattern (typically Gaussian) is to smooth the signal in the sky to suppress fine-scaled detail information of the sky that we know that the telescope will never be able to reconstruct in the first place. However, the sky maps that we simulate to use for testing are already smoothed, such that finer-scale detail of the maps is already removed, this can be seen as:

\[
\vec{d} = A_p(A_s \vec{s}) + \vec{n}_t
\]

Where \( A_p \) is the delta function antenna pattern, applying weight to individual pixels. \( A_s \) is the portion of the antenna pattern that smooths the signal. In actual practice \( A = A_p A_s \). However, for the purpose of our project, \( A = A_p \), while \( (A_s \vec{s}) \) is our effective signal. This estimation allows our matrix \( A \) to be sparse, which is much less computationally heavy.

The mathematical method that we used to reconstruct our maps was used for the COBE mission, named the minimum variance unbiased estimator (Tegmark, 1997), which is:

\[
\hat{s} = (A^\dagger N^{-1} A)^{-1} A^\dagger N^{-1} \vec{d}
\]

Where the noise is contained in the co-variance matrix \( N \): 

\[
\langle n_t n_{t'} \rangle = N_{tt'}
\]

(a) First I will prove that this estimator is unbiased:

\[
\langle \hat{s} \rangle = \langle (A^\dagger N^{-1} A)^{-1} A^\dagger N^{-1} \vec{d} \rangle = \langle (A^\dagger N^{-1} A)^{-1} A^\dagger N^{-1} (A_s \vec{s} + \vec{n}) \rangle = \langle (A^\dagger N^{-1} A)^{-1} (A^\dagger A) \vec{s} \rangle + \langle (A^\dagger N^{-1} A)^{-1} A^\dagger N^{-1} \vec{n} \rangle
\]
\[ = \bar{s} + (A^\dagger N^{-1} A)^{-1} A^\dagger N^{-1} (\bar{n}) \]

The ensemble average of the noise is zero, \( \langle \bar{n} \rangle = 0 \), so:

\[ \langle \hat{s} \rangle = \bar{s} \]

(b) Now, let us assume some matrix \( W \) transforms our data into signal, so:

\[ \hat{s} = W \bar{d} \]

Given that estimator is unbiased, we find the following constraint:

\[ \langle \hat{s} \rangle = \langle W \bar{d} \rangle = \langle W (A \bar{s} + \bar{n}) \rangle = \langle W A \bar{s} \rangle + \langle W \bar{n} \rangle \]

\( \langle \bar{n} \rangle = 0 \), so:

\[ W A \bar{s} = I \bar{s} \]

So our constraint, \( C \), is:

\[ C = W A = I \]

(c) Let \( Q \) be our variance, where:

\[ Q = \langle | \hat{s} - \bar{s} |^2 \rangle = \langle | W (A \bar{s} + \bar{n}) - \bar{s} |^2 \rangle = \langle | W \bar{n} |^2 \rangle \]

We want to minimize \( Q \) subject to the constraint \( C \), so:

\[ \frac{\partial Q}{\partial W} = \Lambda \frac{\partial C}{\partial W} \]

Note that these derivatives are actually applied for each element of \( W \) \((W_{i,j})\), and then inserted back into a matrix. Given that all of our elements for our pointing matrix are real, the result is:

\[ 2 W N = \Lambda A^\dagger = \Lambda A^T \]

Where \( W A = I \), so:

\[ 2 W N W^T = \Lambda A^T W^T = \Lambda (W A)^T = \Lambda \]

\[ 2 W N = 2 (W N W^T) A^T \]

\[ W = W N W^T A^T N^{-1} \]

\[ W A = I = (W N W^T) (A^T N^{-1} A) \]

\[ W N W^T = (A^T N^{-1} A)^{-1} \]
\[ W = (A^T N^{-1} A)^{-1} A^T N^{-1} \]

Thus, we reach our original estimator:

\[ \hat{s} = W \tilde{d} = (A^T N^{-1} A)^{-1} A^T N^{-1} \tilde{d} \]

This gives us a map reconstruction at only a single wavelength bin. The bottom figures display the results of the monochromatic map-making. The maps are displayed in Mollweide view, which gives a two dimensional display of the sky as you would observe it from within the sphere. The first map is randomly generated given a power spectrum (which will be discussed later on), in order to simulate a real map of the CMB. The second map shows that when the telescope makes a number of scans that are not several times larger than the resolution of the sky, the individual scans are not as accurate, and the total map reconstruction fails. The third map is a reconstructed map with a sufficient amount of scans, and the final figure is a reconstructed map with significant noise:

Figure 2: Original simulated sky map of CMB containing 3072 pixels
Figure 3: Reconstructed map of CMB with only 2000 scans and no noise added

Figure 4: Reconstructed map of CMB with 100,000 scans from data with no noise added
We would also like to do map-making in spherical harmonic space for analysis that will be utilized later on. In this case, instead of finding the signal $s_p$ for pixel $p$, we find the coefficients $a_{l,m}$ in a spherical harmonic expansion of the signal, such that for a given pixel:

$$
    s_p = \sum_{l,m} a_{l,m} Y_{lm}(r_p)
$$

Where $Y_{lm}$ is a spherical harmonic function corresponding to a specific $(l,m)$ value (Bunn, 2018). Each $Y_{lm}$ represents a wave on the sphere, where $l$ is its frequency. We can describe the transformation between pixel and harmonic space with a matrix, $\mathbf{Y}$:

$$
    \mathbf{s} = \mathbf{Y}\mathbf{a}
$$

Here, the vector $\mathbf{a}$ contains all of the $a_{l,m}$ coefficients. The resulting new estimator would be:

$$
    \hat{\mathbf{a}} = (\mathbf{Y}^T \mathbf{N}^{-1} \mathbf{A} \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{A}^T \mathbf{d}
$$

Unfortunately, the matrix $\mathbf{Y}$ has complex elements, so it cannot be simply transposed such that this estimator could be further simplified. It is also not sparse like the $\mathbf{A}$ matrix. However, we are able to simulate the application of this matrix using Python’s package named HealPix, which converts between a set of $a_{lm}$ coefficients and a set of pixels for a sky map. Furthermore, the inverse matrix operation that can be seen in both the original map-making estimator as well as the spherical harmonic map-making estimator is solved using the conjugate gradient method, taking advantage of the sparsity of $\mathbf{A}$ and the efficient algorithms for applying $\mathbf{Y}$ to a given vector.
To be clear, we reconstruct maps of the sky that we originally simulate. We use another computational method from Healpix, which takes in a power spectrum and an acceptable number of pixels, and it will create a map with intensity distributions relative to that power spectrum. A power spectrum is the mean-square value of the spherical harmonic coefficients, such that:

\[ C_l = \langle |a_{lm}|^2 \rangle \]

It indicates the amplitude of the waves on the sky as a function of frequency \( l \). The power spectrum we utilize is best fit to the data from the Planck satellite and has been smoothed. For a pixel count of 3072, the power spectrum is shown below:

![Power Spectrum, nPix = 3072](image)

Figure 6: Smoothed Power Spectrum of CMB

### 3 Polychromatic Case

#### 3.1 Eigenvectors/ Eigenvalues

The monochromatic mapmaking case has been long conquered and understood; however, now we are attempting to reconstruct information of multiple wavelength bins from just a single set of scans. In this case, we now consider our signal, either \( s_p \) or \( a_{lm} \), as a function of frequency, \( \nu \). As described earlier, our antenna pattern is also frequency dependent:
\[ d_t = \sum_p \sum_f A(R_t, \nu_f) s_p(\nu_f) + n_t \]

Where \( f \) indicates the discretized frequency bin. Frequency is really continuous so this should be an integral, but we approximate it as a sum. When running our computations to reconstruct maps, we found that 12 wavelength bins for our bandwidth of \((1.5 - 2.5)\text{mm}\) was an appropriate number of discrete units to approximate a continuous case. Specifically, comparing the case of 24 wavelength bins to 12, we found that the uncertainties in our reconstructions remained reasonably constant.

For QUBIC, the distance between its local Gaussian peaks and the center of the beam pattern is proportional to the frequency considered. The figure below shows a two dimensional cross-sectional cut of a Gaussian peak from QUBIC’s antenna pattern.

The first thing to note about reconstructing 12 maps of the sky from one set of scans is that, using our map-making method described up to this point, each of the 12 maps will be terribly inaccurate, with high uncertainty for each pixel/ spatial wave. Therefore, the first step in this new process is to determine what we are able to measure well. There are certain combination of wavelength bins that yield information with low uncertainty. These combinations are the eigenvectors of something named the inverse covariance matrix, which is defined below. To find the best combination of wavelengths, I begin with our estimator:
\[ \hat{s} = (A^TN^{-1}A)^{-1}A^TN^{-1}\vec{d} \]

\[ = (A^TN^{-1}A)^{-1}A^TN^{-1}(A\hat{s} + \vec{n}) \]

\[ = (A^TN^{-1}A)^{-1}(A^TN^{-1}A)\hat{s} + (A^TN^{-1}A)^{-1}A^TN^{-1}\vec{n} \]

\[ = \hat{s} + (A^TN^{-1}A)^{-1}A^TN^{-1}\vec{n} \]

Now I find a set of constants \((v_0, v_1, v_2, v_3...v_{12}) = \vec{v}\) such that \(\vec{v} \cdot \hat{s}\) has low uncertainty:

\[ \vec{v} \cdot \hat{s} = (\vec{v} \cdot \hat{s}) + (\vec{v} \cdot (A^TN^{-1}A)^{-1}A^TN^{-1}\vec{n}) \]

Here, the first term is all the information that we want to know. The second term is the error related to the linear combination of wavelengths. Therefore we want to minimize this error, where:

\[ \sigma^2_v = \langle |\vec{v} \cdot (A^TN^{-1}A)^{-1}A^TN^{-1}\vec{n}|^2 \rangle \]

\[ = \langle |\vec{v}^T(A^TN^{-1}A)^{-1}A^TN^{-1}\vec{n}\vec{n}^T N^{-1}A(A^TN^{-1}A)^{-1}\vec{v}| \rangle \]

Where \(\langle \vec{n}\vec{n}^T \rangle = N\), it simplifies to:

\[ = \vec{v}^T(A^TN^{-1}A)^{-1}\vec{v} \]

We call \(M = A^TN^{-1}A\) the inverse covariance matrix, which is a product of our antenna pattern and the noise of our system.

We want to implement the constraint \(|\vec{v}|^2 = 1\). Then, we want to minimize the variance of our linear combination of wavelengths \(\sigma^2_v\) subject to that constraint:

\[ \frac{\partial\sigma^2_v}{\partial\vec{v}} = \lambda \frac{\partial|\vec{v}|^2}{\partial\vec{v}} \]

\[ 2M\vec{v} = 2\lambda\vec{v} \]

\[ M\vec{v} = \lambda\vec{v} \]

This is an incredibly important result, as it states that the best linear combination of our 12 maps at different wavelength bins are the eigenvectors from the inverse covariance matrix, \(M\). The uncertainty of each eigenvector can be found with a short derivation. Beginning with the last equation above, it can be rewritten as:
\[ M^{-1} \vec{v} = \frac{1}{\lambda} \vec{v} \]

\[ \vec{v}^T M^{-1} \vec{v} = \vec{v}^T \left( \frac{1}{\lambda} \right) \vec{v} \]

\[ \sigma_v^2 = \frac{1}{\lambda} \]

\[ \lambda = \frac{1}{\sigma_v^2} \]

Therefore, the eigenvector with the highest eigenvalue contains the least amount of variance and thus the highest amount of information.

For all-sky experiments, it is much easier to produce eigenvectors from the inverse covariance matrix in spherical harmonic space, such that

\[ \tilde{M} = Y^T A^T N^{-1} A Y \]

This matrix is block diagonal, with a 12 × 12 block that corresponds to each \((l, m)\). Moreover, the blocks depend only on \(l\), not \(m\). This is the same \(l\) from the power spectrum that indicates the frequency of that spatial waveform. This is tremendously useful, because we are able to derive good linear combinations of wavelengths for each \(l\), given by a set of 12 eigenvectors. We are able to derive these eigenvectors for one \((l, m)\) independent of all other \((l, m)\) values; whereas this is impossible in pixel space, because the covariance of one pixel is dependent on all of the remaining pixels. The best three eigenvectors are plotted over 12 wavelength bins spanning between (1.5 – 2.5) mm for an arbitrary \(l\) value below:
Figure 8: Three best eigenvectors with least uncertainty for $l = 18$

The best eigenvector (green) is always approximately constant. This dotted with twelve maps of the sky results in roughly a total intensity for the spatial wave at that $l$ value. The second best eigenvector (orange) is generally a difference between short and long wavelengths - that is, an overall "color" within the bandwidth. We will utilize these eigensystems for polychromatic map reconstruction farther ahead.

3.2 Information Analysis

In the previous section, we produced a different set of eigenvectors for each value of $l$. However, the final goal of this research project is to reconstruct maps using eigenvectors in pixel space instead of harmonic space. Reconstructing full sky maps in harmonic space is completely acceptable; however, the problem is that we want to be able to reconstruct partial sky maps, and then analyze our ability to do so. Therefore, instead of retrieving a set of 12 eigenvectors for each $l$ from $\tilde{M}$, we want to get a set of 12 eigenvectors that are not strictly optimal for any $l$, but that contain most of the information for all $l$. To do so, we sum all of the blocks from the block diagonal $\tilde{M}$ matrix, and extract a set of eigenvectors from this final $12 \times 12$ block. We want to quantify the information that can be reconstructed as:
\[ I = \frac{r_j}{r_j + 1} \]

Here, \( j \) identifies 1 out of 12 eigenvectors for a particular \((l, m)\) spherical harmonic index. \( r_j \) is identified as the squared signal-to-noise ratio for a given mode \( j \). This fraction is chosen because when signal-to-noise ratio is very high, we should get a value of 1, which means this point of information can be measured well. If the signal-to-noise ratio is small, we should get a value of zero, so this point of information cannot be measured well. \( r_j \) is defined as:

\[ r_j = \frac{\langle s_j^2 \rangle}{\sigma_j^2} \]

By definition the magnitude of the angular power spectrum at a given \( l \) is the expected value of signal squared. As derived earlier, the uncertainty squared of the signal for a given eigenvector is the inverse of its eigenvalue, so:

\[ r_j = \lambda_j C_l \]

Thus, the total information possible from all eigenvectors over all \( l \) is:

\[ I_{\text{all}} = \sum_l \sum_n \lambda_{l,m,n} C_l \]

Where \( n \) is an index for the eigenvalue corresponding to 1 of our 12 eigenvectors for a given \((l, m)\). Note that for each \( l \) there are \((2l + 1)\) copies of the same set of eigenvectors and eigenvalues corresponding to different values of the index \( m \). Thus, total information reconstructed is:

\[ I_{\text{all}} = \sum_l \sum_n (2l + 1) \lambda_{l,n} C_l \]

We are trying to determine if it is possible to reconstruct nearly all of the information given by the last equation with our general set of eigenvectors; and if so, how many are sufficient. The formula for the information reconstructed from the general eigenvectors would be:

\[ I_{\text{General}} = \sum_l \sum_n (2l + 1) \lambda_n(\text{General}) C_l \]

Where \( k \leq 12 \). Computationally speaking, the lower number of eigenvectors needed \((k)\), the better.

Below we plot the reconstructed information at each \( l \), as well as the cumulative information. The "All" curves plot \( I_{\text{all}} \), while the other curves plot \( I_{\text{General}} \) for different values of \( k \). The following shows the success story of our research:
Figure 9: Reconstructed info for each $l$ for map with 768 pixels

Figure 10: Cumulative reconstructed info for each $l$ with 768 pixels
Figure 11: Reconstructed info for each $l$ for map with 3072 pixels

Figure 12: Cumulative reconstructed info for each $l$ with 3072 pixels
Figure 13: Reconstructed info for each $l$ for map with 12288 pixels

Figure 14: Cumulative reconstructed info for each $l$ with 12288 pixels
One will notice that as the resolution of the map increases, the number of eigenvectors required to reconstruct all of the possible information increases. In all examples, our method for reconstructing maps in pixel space is shown to be possible, as it only requires a handful of eigenvectors to accomplish.

The next two graphs show the information plots when there is an offset of the beam pattern. The distance between peaks ($d_{\text{peak}}$, shown in the contour map in figure (1)) is constant for a given wavelength, but these localized peaks are able to be offset, in which the center gaussian peak is translated away from the peak of the large, virtual gaussian envelope. An offset can occur in two dimensions, such that if a peak is translated a distance $d_{\text{peak}}$ horizontally, a full revolution would be made and the antenna pattern would remain the same. Thus, the largest notable offset in the antenna pattern would be $\frac{d_{\text{peak}}}{2}$ in the horizontal and vertical directions. The plots below consider that case.

![Figure 15: Reconstructed info for each $l$ for map with 3072 pixels. Antenna pattern is offset by $(x, y) = (50\% d_{\text{peak}}, 50\% d_{\text{peak}})$.](image)
Figure 16: Cumulative reconstructed info for each $l$ with 3072 pixels. Antenna pattern is offset by $(x, y) = (50\% d_{\text{peak}}, 50\% d_{\text{peak}})$.

Although this is the most extreme offset, it is clear that a large offset in the antenna pattern makes the signal much more difficult to reconstruct.

### 3.3 Map Reconstruction

Polychromatic map reconstruction using our general eigenvectors is possible in pixel and harmonic space; however, we must perform the reconstruction in pixel space for partial sky data. We do this by creating a matrix $U$. This matrix is large and block diagonal, in which each block consists of the $k$ best general eigenvectors that are required to reconstruct all of the information, inserted as its columns. As discussed before, one eigenvector gives a linear combination of 12 different maps (each map corresponding to a different wavelength bin). This means that for each pixel on the map, one eigenvector yields one number. Therefore, if $k$ different general eigenvectors are applied to twelve different maps, we will receive $k$ maps, that we would say exist in eigenspace. This can be shown by:

$$\vec{s} = U\hat{s}$$

Where $\vec{s}$ contains our original 12 maps in pixel space. $\hat{s}$ is our $k$ maps in eigenspace. Thus, our time-ordered data equation would be:

$$\vec{d} = (A U)\hat{s} + \vec{n}$$
Therefore, our new map-making estimator would become:

$$\hat{s} = (U^T A^T N^{-1} A U)^{-1} U^T A^T N^{-1} \tilde{d}$$

For the case of a resolution of 3072 pixels, we see in figure (11) that the three best eigenvectors reproduce all of the information. Thus, we can transform 12 maps in pixel space to 3 maps in eigenspace. As an example of our results, the first figure below shows the one of three map in eigenspace that corresponds to the best eigenvector (which has the lowest uncertainty). The 3 simulated maps are translated into time-ordered data by the equation above. That data is taken and the estimator above is used to reconstruct the same three maps, where figure (18) is the reconstruction of the map in figure (17). The third figure below shows the map reconstructed in the case where our telescope only points at the top hemisphere of the sky.

Figure 17

Figure 18
4 Conclusion and Future Works

In a QUBIC-like experiment, the time-ordered data will contain multiple independent color combinations that we now see can be reconstructed with reasonable signal-to-noise. At fairly high resolution, we have successfully developed a procedure that reconstructs multiple maps in a way that contains virtually all available information. This remains true even when we reconstruct only part of the sky. We can see from all of the information plots that the best eigenvector (the one with the highest eigenvalue) yields majority of the possible information able to be reconstructed.

One next step in our research includes applying the map reconstruction procedure to realistic data, such as CMB maps with dust and synchotron contamination. The most difficult next step would be quantifying the information reconstructed in the partial sky case into the same plots we see in figures (9)-(16). This is an issue because we can no longer simulate our signal squared with the power spectrum, as this approximation is only valid in the full sky case.

5 References


