ILC Analysis of BICEP2 and Keck Array CMB Polarization Data through the 2015 Observing Season

A dissertation submitted to the
Graduate School of the University of Cincinnati
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in the Department of Physics
of the College of Arts and Sciences

by

Emma Hand

June 2023

supervised by

Dr. Colin Bischoff

Committee Member: Dr. Kay Kinoshita, Physics

Committee Member: Dr. Rohana Wijewardhana, Physics

Committee Member: Dr. Adam Aurisano, Physics

Abstract

The cosmic microwave background (CMB) is a snapshot of the universe at recombination, the moment when the universe became transparent. Understanding the CMB could allow us to constrain or rule out aspects of inflation theory, which suggests that the Universe underwent a period of rapid expansion mere moments after the Big Bang. Specifically, we hope to detect primordial gravitational waves (PGW), an as yet unobserved phenomenon predicted by many inflationary models. BICEP/Keck is an experiment based at the South Pole with telescopes that are specified to observe B-mode polarization patterns caused by PGW at microwave wavelengths. The tensor-to-scalar ratio r is a parameter which if constrained could provide indirect evidence of PGWs. Presented is the work I have done in an internal linear combination (ILC) component separation method to separate the CMB signal from galactic foregrounds while minimizing noise, and the likelihood analysis I performed with the ILC results in an effort to constrain r.

Acknowledgements

This work would not have been possible were it not for the many people in my life, both academically and personally, who were with me through the entire process. First I would like to thank my advisor, Colin, who was always happy to answer the thousands of questions I had along the way, offering valuable insights and guidance throughout the process. I'd like to thank Caterina, who was extremely helpful during my on-boarding period, helping me set up my workspace and providing materials and guidance to help me get me familiar with the work. I'd like to thank the Physics department as a whole, the many professors of which gave me the foundational education to survive graduate school. To Rachel, who was with me even from across the country, working on her own thesis, who was a friend I could always count on.

I want to thank my family and friends for their unconditional love and support through my years in graduate school. To my mom and siblings, thank you for all of the laughter which was absolutely necessary to push me through. To my husband Jacob, whose unwavering support was always offered without question. And finally my wonderful daughters, Zo and na, for always forgiving the amount of time I had to step away for work in hopes of providing a better future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cosmology

The Universe is a vast, empty space that is continuously expanding. The study of the beginning of the Universe and how it evolved to its current state is called cosmology. Many decades of ground and space based observations have given us the standard cosmological model ΛCDM which tells us that the makeup of the Universe is cold dark matter (CDM), baryonic matter, and dark energy (associated with a cosmological constant denoted by Λ, which is a favored specific model of dark energy) [1], which drives the acceleration of the expansion of the Universe. It tells us that the Universe began 13.7 billion years ago in the Big Bang and has been expanding ever since. It does not, however, explain the flatness of the Universe, the horizon problem, or how large scale structure was able to occur. Inflation theory postulates that the Universe rapidly expanded during a period of approximately 10^{-35} seconds after the Big Bang, and can address these shortcomings of ΛCDM. The cosmic microwave background (CMB) is a remnant radiation field that gives us a snapshot of the Universe just before recombination, the study of which can strengthen the argument for Inflation theory. For this reason the BICEP/Keck collaboration specializes in looking

for *B*-mode patterns in the CMB polarization. This chapter will delve into the standard cosmological model and inflation, and provide a deeper picture as to why the CMB is central to understanding the early Universe.

1.1.1 The Standard Model of Cosmology

The standard model of cosmology assumes general relativity and invokes the cosmological principle that the universe at large scales is homogeneous and isotropic. The following is a brief derivation of the cosmological model beginning with the Einstein field equations.

At large scales, the behavior of the Universe can be described with the Einstein field equations [2]:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1.1}$$

where R and $R_{\mu\nu}$ are the Ricci curvature scalar and tensor, Λ is the cosmological constant, $g_{\mu\nu}$ is the metric tensor which describes the geometry of space-time, G is Newton's constant and $T_{\mu\nu}$ is the stress-energy tensor which describes density and flux of momentum and energy in space-time. Essentially, equation (1.1) relates the dynamics between the geometry and energy of space-time, and solutions to this equation require some knowledge of the contents and assumptions of the space-time metric. Assuming the cosmological principle, we can use the Friedmann-Robertson-Walker (FRW) metric ds^2 to describe space-time:

$$ds^2 = c^2 dt^2 - a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 d\theta + r^2 sin^2\theta d\phi^2 \right) \tag{1.2}$$

where t is the temporal coordinate, (r, θ, ϕ) are the spatial comoving coordinates, and k is spatial curvature ($k = 0$ for a flat universe). The scale factor $a(t)$ describes the expansion rate of the Universe and is commonly set to 1 for the present time, and was smaller in earlier
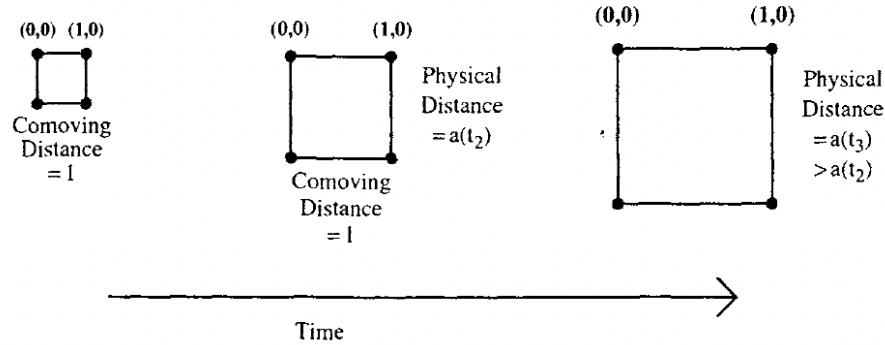
Figure 1.1: Simple diagram showing expansion of space. The comoving distance between points remains the same over time, however the physical distance between the two points increases. The relationship between these two distances relies on the scale factor a. Figure from Ref. [1].

times than it is today. The diagram in figure 1.1 helps give an intuitive sense of the scale factor a. We can think of a square connecting points on a grid which grows over time such that the distance between the points increases, but the points on the grid maintain their values. The comoving distance between two points is the difference in the coordinates at which the points rest, and with time this difference remains the same. However the physical distance between the points grows, and this distance is proportional to the comoving distance times the scale factor a.

For sufficiently large scales, the contents of the universe can be described as a perfect fluid, and the stress-energy tensor can be described in terms of energy density ρ and pressure p. This simplification as well as the FRW metric in equation (1.2) allow us to solve the Einstein equations to get the Friedmann Equations [3]:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2} \tag{1.3}$$

$$\frac{\ddot{a}}{a} = \frac{\Lambda c^2}{3} - \frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) \tag{1.4}$$

where H is the Hubble parameter which measures the rate at which the scale factor changes.

3
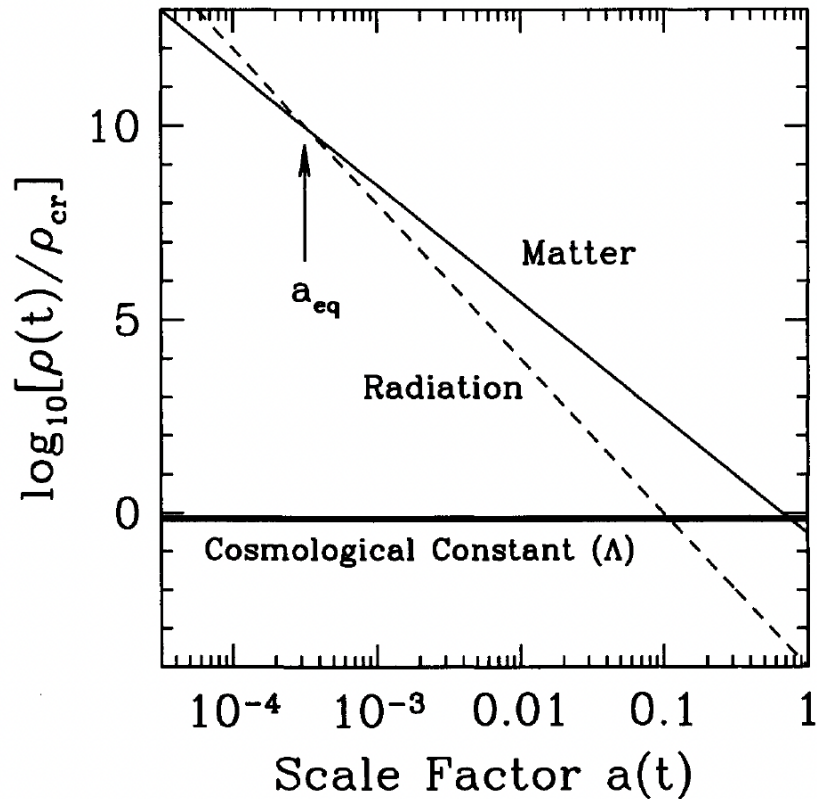
Figure 1.2: Energy density (in units of today's critical density) versus a for the different energy components of a flat universe. Solid black is nonrelativistic matter, dashed is radiation, and bold solid is a cosmological constant Λ. a_eq is the point at which radiation and matter were equal. Figure from Ref. [1].

It is necessary to know how the energy density evolves with time in order to be able to use Einstein's equations. This requires an understanding of the contents of the universe, since each distinct component scales differently with a. Matter density scales inversely with its volume, therefore $\rho_m \propto a^{-3}$, while the wavelength of radiation stretches with expansion and has a more rapid scaling of $\rho_r \propto a^{-4}$. This relationship comes from the fact that the energy of a photon is inversely proportional to its wavelength $(1/a)$, so the energy density then would be its energy multiplied by the number density (which is inversely proportional to volume, or $1/a^3$). Finally, dark energy appears to have a relatively constant density over time and $\rho_\Lambda \propto a^0$. Figure 1.2 shows the relationship between energy density of the different components (for a flat universe) and the scale factor.

With these relationships we can find what dominated the expansion of the early Universe. When the scale factor a was small radiation dominated the expansion of the Universe, and because the energy density faded more quickly than matter density, matter eventually dominated the acceleration. As matter density faded an era of dark energy density domination ensued, an era which continues to this day.

There is a critical density for which the Universe is spatially flat, $\rho_{cr} = \frac{3H^2}{8\pi G}$. This density separates closed ($\rho > \rho_{cr}$) and open universes ($\rho < \rho_{cr}$). Using a ratio between energy density and this critical density, $\Omega = \frac{\rho}{\rho_{cr}}$, equation (1.3) can express the relationship between the contents of the Universe and the dynamics of the scale factor as:

$$\frac{H^2}{H_0^2} = \Omega_{0,r}a^{-4} + \Omega_{0,m}a^{-3} + \Omega_{0,k}a^-2 + \Omega_{0,\Lambda}a^0 \tag{1.5}$$

where H_0 is today's Hubble constant and each of $\Omega_{0,r}, \Omega_{0,m}$ and $\Omega_{0,\Lambda}$ describe the current energy densities. $\Omega_{0,m}$ can be further broken down into baryonic matter ($\Omega_{0,b}$) and dark matter ($\Omega_{0,d}$) (evidence for which can be found imprinted in the CMB temperature and polarization power spectra [4]), which are stable particles with negligible electromagnetic interactions [5]. $\Omega_{0,k} = 1 - \Omega_0$ is the current spatial curvature density (where Ω_0 is the sum the energy densities described above). The Planck Collaboration's most recent results [6] find that the Universe is broken down into 31% matter (or 5% baryonic matter and 26% cold dark matter), 69% dark energy (Λ) and negligible (0.005%) radiation, and a spatial curvature density ≈ 0.001, meaning the universe can be approximately modeled as flat with a spatial curvature parameter $k = 0$.

1.1.2   Inflation

While ΛCDM excellently describes the evolution of the Universe as we see it today, it does not describe the initial conditions that caused the current state of the Universe, and still has

some open questions:

- **The Horizon Problem**: The temperature of the sky that we observe is uniform on the order of 1 part in $10,000$ in all points of the sky. According to ΛCDM alone, the maximum distance two particles of light in causal contact could have traveled is approximately a $2°$ separation on the sky [7]. Why then do we observe photons in all directions of the sky measure to be the same temperature?

- **The Flatness Problem**: As mentioned above, the Universe is observed to be spatially flat. Given the infinite values that Ω_k could take for an open ($\Omega_k < 0$) or closed ($\Omega_k > 0$) universe, the existence a flat universe requires extremely fine tuned conditions.

- **The Source of Initial Perturbations**: While ΛCDM provides explanation for the existence of large scale structure, it does not explain the source of the initial perturbations from which the large scale structures emerge.

Inflation is an extension to the ΛCDM model first proposed over 40 years ago [8] modeling the very early stages of the universe and can resolve the above problems. It postulates that the space-time exponentially expanded by ≈ 60 e-folds in the span of 10^{-35} seconds after the Big Bang before continuing to standard expansion. This extreme growth in such a small amount of time can explain how it was possible for particles on opposite sides of the sky to have been in thermal equilibrium, explaining why the universe is isotropic and homogeneous and solving the horizon problem. A flat universe can be explained in a similar way. At sufficiently small scales space-time is locally flat, therefore after inflation the flat curvature is preserved. The large scale structure is sourced from quantum fluctuations that are expanded to astrophysical scales with inflation [9].

1.1.3 Single Field Slow Roll Inflation

Inflation is a relatively simple theory, requiring an exponential expansion for a sufficient period of time in order to solve the above questions. The dominant class of inflation models, Single Field Slow Roll, proposes a scalar inflation field Φ which dominates the energy density with energy potential $V(\phi)$. The pressure and density for such a field can be written as:

$$p = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{1.6}$$

$$\rho = \frac{1}{2}\dot{\phi}^2 - V(\phi) \tag{1.7}$$

Taking the time derivative of equation (1.3) along with the definitions of pressure and density above, we arrive at the equation of motion for inflation field:

$$\ddot{\phi} + 3H\dot{\phi} - \frac{dV}{d\phi} = 0 \tag{1.8}$$

where $3H\dot{\phi}$ is the dampening term, also known as Hubble friction. An accelerated expansion requires that $\ddot{a} > 0$, which from equation (1.4) implies that $\rho + 3p < 0$ (where we are setting the speed of light $c = 1$). This condition tell us that $p < 0$ (because density is always positive) and from equation (1.6) we see that this is satisfied for $\dot{\phi} \ll V(\phi)$, meaning that the inflation field must evolve slowly. For this to be the case the kinetic term cannot overwhelm the potential one, therefore from equation (1.8) we get the condition $\ddot{\phi} \ll 3H\dot{\phi}$.

These two conditions are often written as the slow roll parameters [8]

$$\epsilon = 3\frac{\dot{\phi}^2/2}{V + \dot{\phi}^2/2} \simeq \frac{M_{Pl}^2}{2}\left(\frac{V'}{V}\right)^2 \ll 1 \tag{1.9}$$

$$\eta = -2\frac{\dot{H}}{H^2} - \frac{\dot{\epsilon}}{2H\epsilon} \simeq \frac{M_{Pl}^2}{2}\left(\frac{V''}{V}\right)^2 \ll 1 \tag{1.10}$$

where M_{Pl} is the reduced Planck mass $8\pi G^{-1/2}$ and V' and V'' are the first and second derivatives of V with respect to ϕ.

During the period of inflation the inflaton energy density is dominant. Quantum mechanical fluctuations in the inflaton field become very important as they induce scalar perturbations in the space time metric that take form as matter density fluctuations and tensor perturbations which correspond to primordial gravitational waves (PGW).

The scalar perturbations can be parameterized by a power law in Fourier space

$$P_s(k) = \Delta_s \left(\frac{k}{k_{s*}} \right)^{n_s-1} \tag{1.11}$$

where k is the wave number and k_{s*} is the pivot scale (wave number) for which amplitude Δ_s has been specified, and n_s is the scalar spectral index describing how the scalar fluctuations vary with scale. The generic models of inflation state that the scalar perturbations are scale-invariant corresponding to $n_s - 1 = 0$, however this is for a special space time and realistically n_s is expected to deviate slightly from 1 [8] (Planck measures a value of ≈ 0.96 [6]).

The tensor perturbations occur due to the PGW stretching and compressing space, and can similarly be described with a power law

$$P_t k = \Delta_t \left(\frac{k}{k_{t*}} \right)^{n_t} \tag{1.12}$$

where again k_{t*} is the pivot scale for defined amplitude Δ_t, and $n_t \approx 0$ is the tensor spectral index.

The tensor-to-scalar ratio r is what gives PGW amplitude and is defined as

$$r = \frac{\Delta_t}{\Delta_s} \tag{1.13}$$

The detection of this ratio would provide an amplitude for PGW and strong evidence for Inflation, and give information on the energy scale of Inflation which is directly related to the strength of the gravitational waves. The direct detection of PGW is not currently possible due to the low signal amplitude, however looking for certain imprints on the CMB provides a way in which this parameter can be measured indirectly.

1.2 The Cosmic Microwave Background

The CMB is the earliest remnant of the universe that we can directly observe. The photons from the CMB last scattered 380,000 years after the Big Bang, when the universe was still opaque. At the time when the CMB photons decoupled their temperatures were several thousand Kelvin. Due to the continuing expansion of the universe, the photon wavelengths have been stretched from the visible spectrum to microwave wavelengths, a process which cooled the photons. Today the CMB closely follows the spectrum of a 2.7 Kelvin blackbody with temperature fluctuations on the order of 10^{-5} Kelvin [5]. These fluctuations are due to density fluctuations which eventually evolve into large scale structure [9]. Figure 1.3 shows the CMB blackbody spectrum.

1.2.1 CMB Anisotropies

The small fluctuations in the CMB temperature are where we are able to glean the wealth of information that the statistics of the CMB provide. The CMB can be thought of as a snapshot of the early universe at the time of recombination, as it is the moment that photons were able to escape the primordial plasma that described the early universe.

During the time before recombination, there were constant reactions between the gravitational pull of the matter and the counteractive photon pressure, causing oscillations in the baryonic fluid which lead to density perturbations corresponding to the hot and cold
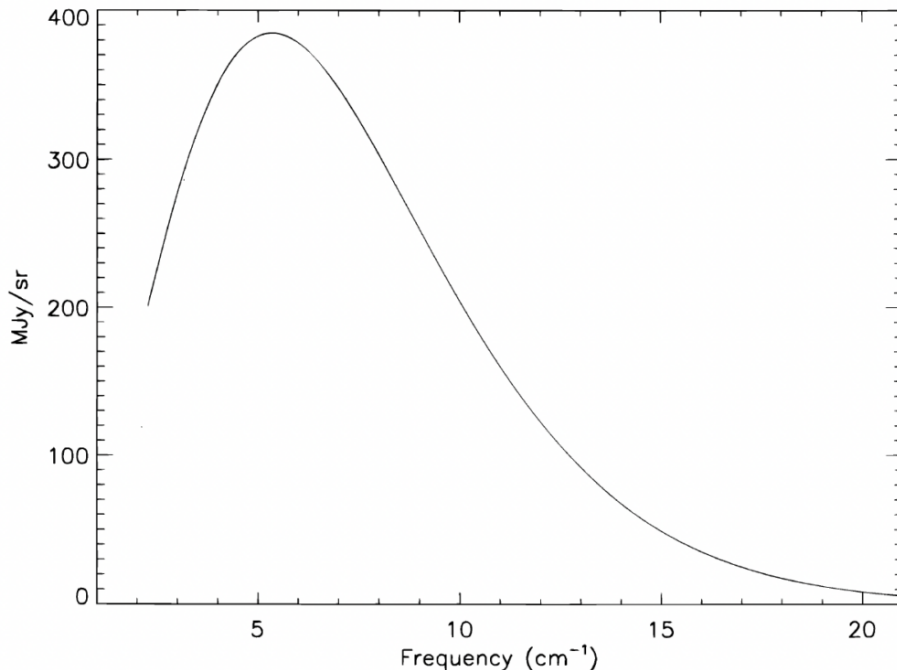
Figure 1.3: CMB blackbody spectrum as measured by the FIRAS instrument of the COBE experiment, with error bars a fraction of the width of the best fit line. [10]

regions of the patterns imprinted on the CMB temperature. At the time of recombination we see that some modes were at maximum compression for the first time while others were at their first rarifaction, which caused the largest hot and cold spots. Other modes were at their second contraction/rarifaction, leading to hot and cold spots half the angular size of the maximum modes, and so on.

The temperature anisotropies (ΔT) at some point (θ, ϕ) in the sky can be modeled in spherical harmonics with amplitude $a_{\ell,m}$:

$$\Delta T(\theta, \phi) = \sum_{\ell=1}^{\inf} \sum_{m=-\ell}^{\ell} a_{\ell,m} Y_{\ell,m}(\theta\phi) \tag{1.14}$$

where θ and ϕ are the angles describing the celestial sphere, ℓ corresponds to angular size $(\theta \approx 180°/\ell)$ and m is the number of modes at scale ℓ. These anisotropies are shown in figure 1.4 as measured by Planck. While we are not able to predict the actual values of $a_{\ell,m}$, we can predict a Gaussian distribution, and then extract the cosmological information from
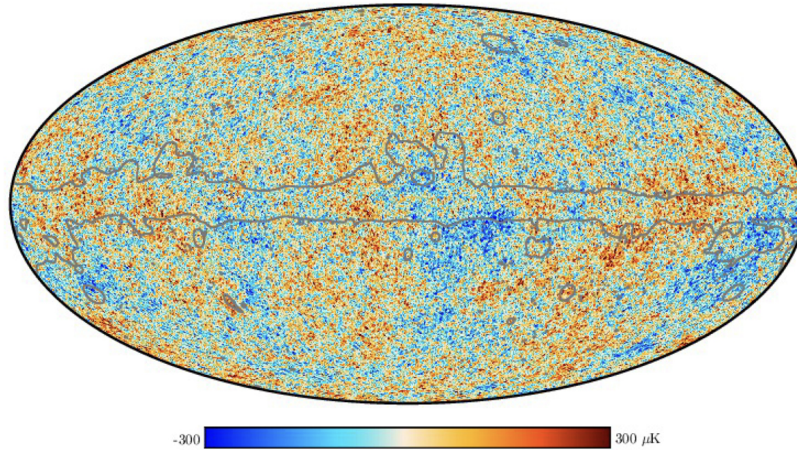
Figure 1.4: Planck CMB temperature map with mean temperature subtracted as measured by SMICA [11]. This map has had residual foregrounds removed, and a mask applied to remove the galactic plane, as endicated by the grey lines towards the center of the map.

the CMB angular power spectrum (APS) which is the variance on the amplitudes:

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{m=\ell} a_{\ell,m} a_{\ell,m}^* \tag{1.15}$$

It is common in cosmology to use the scaled power spectrum $D_\ell = \ell(\ell+1)C_\ell/2\pi$, so that ℓ modes which have not gone through full oscillations are essentially flat.

In practice we can calculate the APS from the map of the CMB by "crossing" the map by itself (auto power spectra), or crossing it with another map (cross power spectra). For example, a cross of a CMB polarization map with a CMB temperature map will yield a cross spectrum, while a cross of a temperature map with itself will yield an auto spectrum.

The variance on a power spectrum is called the cosmic variance, and is given by

$$\frac{\Delta C_\ell}{C_\ell} = \sqrt{\frac{2}{2\ell + 1}} \tag{1.16}$$

This is the theoretical limit on the measurement of the CMB power spectrum coming from the fact that we only have one sky to measure. Figure 1.5 shows the CMB temperature power
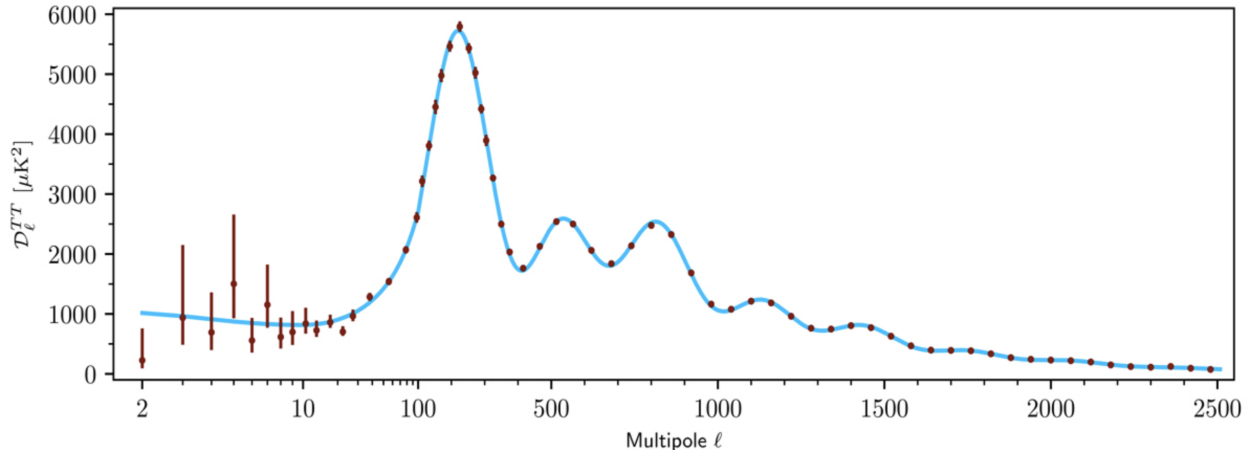
Figure 1.5: CMB temperature power spectrum as measured by Planck [11]. The data points are shown in red, while the blue line shows the ΛCDM best-fit model. Note that the x-axis scale changes from log to linear at $\ell = 100$.

spectrum as measured by Planck [11], where the blue line is the best fit ΛCDM model.

The power spectrum for $\ell < 20$ indicates the initial conditions of the universe, with modes corresponding to Fourier modes at time of decoupling that had wavelengths longer than the horizon scale. The first three peak positions and relative heights tell us that the universe is spatially approximately flat, and give constraints on the baryonic and dark matter [5].

1.2.2   CMB Polarization

The CMB is a polarized field due to Thomson scattering at the time of recombination, in which an incoming unpolarized photon scattering off of a free electron is converted to a photon with polarization perpendicular to the line of sight. Due to the fact that Thomson scattering was happening constantly at all angles before recombination, the final polarization of the escaping photon depends on the local temperature anisotropy in which the photon scattering off of the electron occurs.

For example, photons scattering off of an electron in a local temperature dipole will have no final net polarization, as the polarization aligned with the 180° temperature difference will average out and match the polarization at 90°. Photons scattering off of an electron in

a temperature monopole, a local uniform temperature, will also yield a collection of photons with no final net polarization. A photon scattering off of an electron sitting in a local temperature quadrupole, however, will result in a linearly polarized photon. Because the requirement for a photon from last scattering to be linearly polarized is local termperature quadrupoles, the CMB polarization can be thought of as a map of these temperature quadrupoles on the surface of last scattering.

Stokes Parameters

The polarization of photons traveling in the $\hat{z}$ direction can be described by equation (1.17)

$$E(x,t) = E_x cos(\omega t - \theta_x)$$
$$E(y,t) = E_y cos(\omega t - \theta_y)$$

(1.17)

where E is the electric field with corresponding amplitudes E_x and E_y, ω is the frequency of the wave, t is time, and θ_x and θ_y are the phase of the wave. This electric field can then be broken down into the Stokes' parameters in equation (1.18)

$$I = \langle E_x^2 \rangle + \langle E_y^2 \rangle$$
$$Q = \langle E_x^2 \rangle - \langle E_y^2 \rangle$$
$$U = 2\langle E_x E_y \rangle cos\phi$$
$$V = 2\langle E_x E_y \rangle sin\phi$$

(1.18)

where ϕ is the phase difference between x and y components, and $\langle \cdot \rangle$ is the time average. I is the radiation intensity invariant under rotation of the x and y axes, Q and U are the linearly polarized light in $+$ and $\times$ directions and are coordinate dependent, and V is circularly polarized light. Because Thomson scattering can only produce linear polarization for the CMB, V is 0.

We are interested in the power spectrum of the CMB polarization just as we are the CMB temperature. We can expand the polarization maps in spherical harmonics just like the temperature, however the harmonics must be weighted with spin ± 2 as the two quantities $Q \pm iU$ are spin ± 2 quantities [12].

$$(Q \pm iU)(\hat{n}) = \sum_{\ell,m} {}_{\pm 2}a_{\ell,m\pm 2}Y_{\ell,m}(\hat{n}) \tag{1.19}$$

E and B-mode Polarization

It is convenient to work in a basis that is coordinate independent, and this can be done by considering the rotationaly invariant coefficients and basis functions [12]

$$
\begin{aligned}
a_{\ell,m}^E &= \frac{-\left({}_{+2}a_{\ell,m} + {}_{-2}a_{\ell,m}\right)}{2} & X_{1,\ell,m} &= \frac{{}_{+2}Y_{\ell,m} + {}_{-2}Y_{\ell,m}}{2} \\
a_{\ell,m}^B &= \frac{-i\left({}_{+2}a_{\ell,m} - {}_{-2}a_{\ell,m}\right)}{2} & X_{2,\ell,m} &= \frac{{}_{+2}Y_{\ell,m} - {}_{-2}Y_{\ell,m}}{2}
\end{aligned}
\tag{1.20}
$$

We can completely write the linear combination of the Stokes Q and U into a coordinate independent E and B basis

$$
\begin{aligned}
P(\hat{\mathbf{n}}) \equiv \begin{bmatrix} Q(\hat{\mathbf{n}}) \\ U(\hat{\mathbf{n}}) \end{bmatrix} &= -\sum_{\ell=2}^{\inf}\sum_{m=-\ell}^{\ell} a_{\ell,m}^E \begin{bmatrix} X_{1,\ell,m}(\hat{\mathbf{n}}) \\ -iX_{2,\ell,m}(\hat{\mathbf{n}}) \end{bmatrix} + a_{\ell,m}^B \begin{bmatrix} iX_{2,\ell,m}(\hat{\mathbf{n}}) \\ -X_{1,\ell,m}(\hat{\mathbf{n}}) \end{bmatrix} \\
&= -\sum_{\ell=2}^{\inf}\sum_{m=-\ell}^{\ell} a_{\ell,m}^E Y_{\ell,m}^E(\hat{\mathbf{n}}) + a_{\ell,m}^B Y_{\ell,m}^B(\hat{\mathbf{n}})
\end{aligned}
\tag{1.21}
$$

where $Y_{\ell,m}^E$ and $Y_{\ell,m}^B$ are the basis functions for the linear polarization P field on a sphere, and E-mode $Y_{\ell,m}^E$ and B-mode $Y_{\ell,m}^B$ are patterns in the CMB polarization analogous to the curl-free electric and divergence-free magnetic fields. Figure 1.6 shows a cartoon illustrating the production of E and B-mode polarization. Under parity transformation, E-modes remain
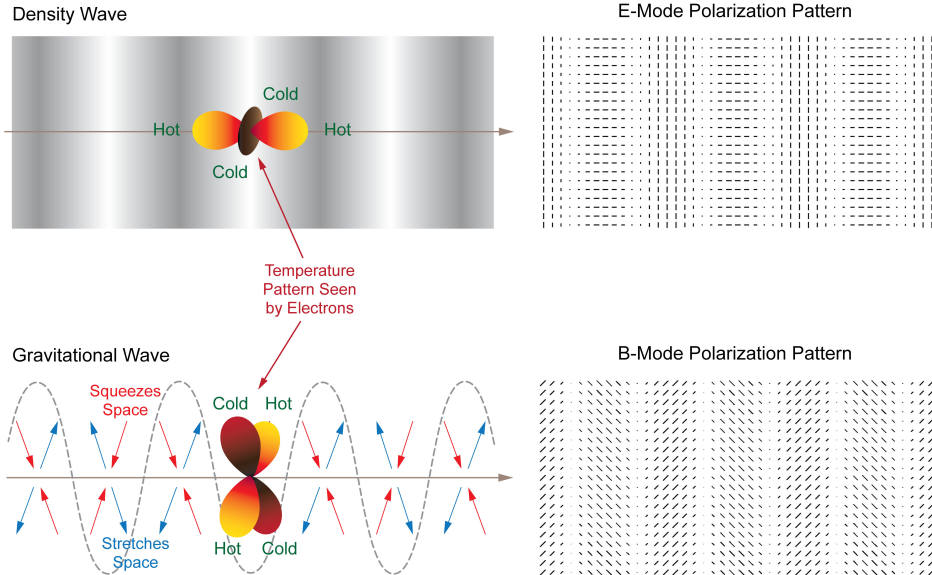
Figure 1.6: A pictoral representation density and gravitational waves producing E-mode and B-mode polarization. An electron through which a single density wave which propagates as shown in the top left panel will see hotter and colder regions in either the parallel or perpendicular direction to the plane of the density wave. This wave will only produce E-mode polarization patterns, shown directly to the right. If now instead a single gravitational wave propagates through the electron in the same direction as the density wave, it can warp the space in a direction perpendicular from its propagation, as shown in the bottom left panel. This type of wave, depending on the orientation of the stretch/squeeze motion, can create either E-mode or B-mode patterns (shown on the bottom right). Diagram curtesy of BICEP/Keck.

the same while B-modes flip signs, allowing for differentiation between their sources. E-modes in the CMB can be sourced from both scalars (temperature, density) and tensors (PGW), while B-modes can only be sourced from tensors.

Similar to the CMB temperature power spectrum, we can obtain the polarization EE and BB power spectra

$$C_\ell^{EE} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} a_{\ell,m}^E a_{\ell,m}^{E*} \tag{1.22}$$

$$C_\ell^{BB} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} a_{\ell,m}^B a_{\ell,m}^{B*} \tag{1.23}$$

15

While in practice we can obtain power spectra between the fields, the only non-zero cross spectrum is the TE power spectrum. This is due to the fact that the CMB power spectra are invariant under parity transformation according to the ΛCDM model, hence the cross between even-parity (E) and odd-parity (B) quantities must be zero.

Primordial Gravitational Waves

PGWs are a phenomenon predicted by inflation which have not yet been directly observed. Immediately after the Big Bang there were perturbations in the gravitational metric, and with inflation these perturbations would have stretched out to become PGWs. These are a standard prediction for many inflationary models and their discovery would be a strong source of evidence favoring inflation [13]. If PGWs exist, they would leave an imprint on the polarization of the CMB, and because PGWs are the only source that could have imprinted B-mode patterns we focus on the study of B-modes to find evidence of PGWs.

E modes have much higher amplitudes than B modes [8]. The amplitude of PGW is set by the parameter r, the ratio between tensor power and scalar power, and is directly proportional to the energy scale of inflation. Inflation predicts a peak in the B-mode power spectrum around $\ell = 80$, or about a two degree scale on the sky.

1.3 Foregrounds

There are other components in the sky that have signal in the same frequency range as the CMB, the brightest of which are the Galactic material foregrounds. The Galaxy emits radiation in the microwave, some of which is polarized and can look like patterns expected in the CMB polarization. Dust can be aligned with the magnetic field of the Milky Way causing B-mode patterns in the signal we observe, and is brightest at high frequencies. Free electrons spiraling in the magnetic field of the galaxy emit synchrotron radiation and produce B-mode
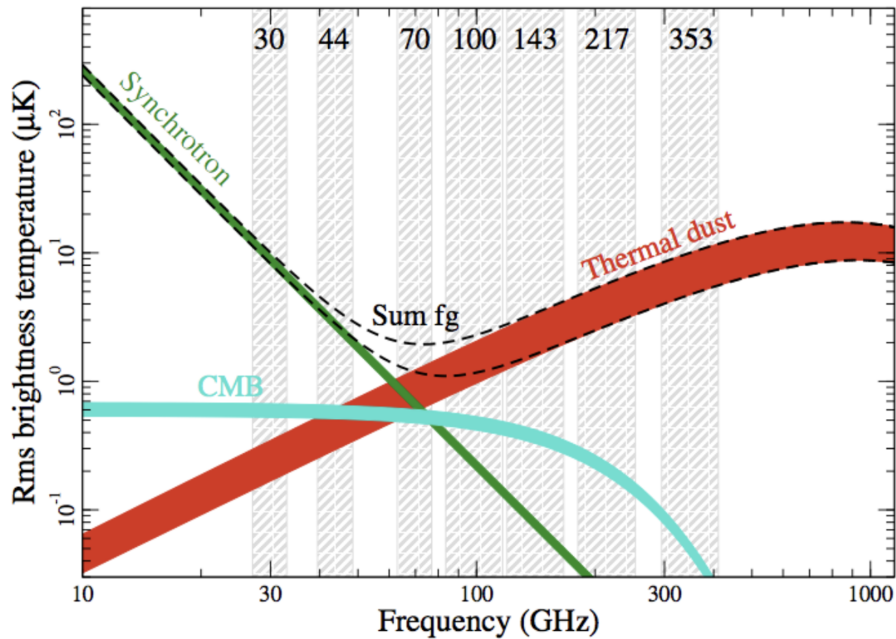
Figure 1.7: Submillimeter sky dust and synchrotron polarization frequency dependence. The grey bands show the frequecy channels for Planck, the green band indicates synchrotron, red indicates dust, and blue is CMB. As this figure is from a specific fraction of sky, we note that the foregrounds are not necessarily larger than the CMB for all small patches of sky. Figure from [11].

patterns brightest at low frequencies [5]. Figure 1.7 shows the spectral characteristics of the foregrounds in polarization with respect to frequency [14].

Dust

The dust in the Milky Way emits radiation at a characteristic temperature of $T_d \approx 20K$, with a frequency dependence that follows a modified blackbody spectrum, shown in equation (1.24).

$$I_d(\nu) \propto \nu^{\beta_d} B(\nu, T_d) \tag{1.24}$$

where ν^{β_d} accounts for the dust emissivity as a function of frequency [15] and $B(\nu, T_d)$ is the Planck function given as

A_d^P
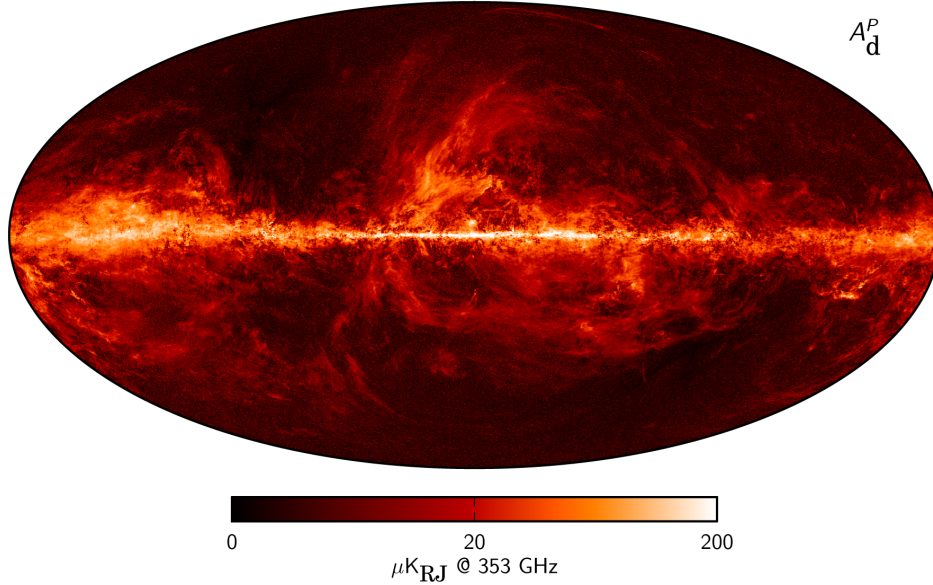
Figure 1.8: Planck dust polarization amplitude map at 353 GHz, where $P = \sqrt{Q^2 + U^2}$ produced from Planck and WMAP data. Image from Ref. [17].

$$B(\nu, T_d) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(\frac{h\nu}{k_B T}) - 1} \tag{1.25}$$

While the science of the dust in the Milky Way is quite complex [16], for the purposes of our work it can be thought of as elongated grains, with their major axis aligning with the magnetic fields of the galaxy. They emit more efficiently along their shorter axis, creating a net polarized emission. Because of the frequency dependence of dust polarized radiation, observing at varying frequencies makes it possible to constrain and remove this signal from CMB polarization signal. The concentration of the dust grains is highest along the Galactic plain, so observing outside of this region allows for minimization of dust interference in CMB data. Figure 1.8 shows the Planck dust polarization amplitude map at 353 GHz.

Synchrotron

Cosmic ray electrons accelerating in the Milky Way emit synchrotron radiation, which is partially polarized perpendicular to the magnetic field [18]. The intensity of the radiation

18

is dependent on both the cosmic ray density (n_e) and the strength of the magnetic field perpendicular to the line of sight $(B_\perp)$. We can write synchrotron intensity for frequency ν (for a cosmic ray distribution which follows the power law $N(E) \propto E^{-s}$ where s is cosmic ray spectrum) as

$$S(\nu) = \epsilon_s(\nu) \int_z n_e B_\perp^{(1-s)/2} dz \qquad (1.26)$$

where the integral is over the direction of line of sight z and $B_\perp = \sqrt{B_x^2 + B_y^2}$ is the magnetic field for the x-y axes being the plane of the sky. $\epsilon_s(\nu)$ is the emissivity term given by

$$\epsilon(\nu) = \epsilon_0^{\beta_s} \qquad (1.27)$$

and

$$\beta_s = -\frac{s+3}{2} \qquad (1.28)$$

We can then model the synchrotron intensity as

$$S(\nu) = S(\nu_0)(\frac{\nu}{\nu_0})^{\beta_s} \qquad (1.29)$$

where spectral index β_s has a best fit $\beta_s \sim -3$ from WMAP and Planck low frequency band observations [19]. Figure 1.9 shows the Planck synchrotron polarization amplitude map at 30 GHz.

Gravitational Lensing

The CMB photons that reach us have traversed the radius of the universe, which has a distribution of large gravitational sources. This distribution creates a lensing potential ϕ
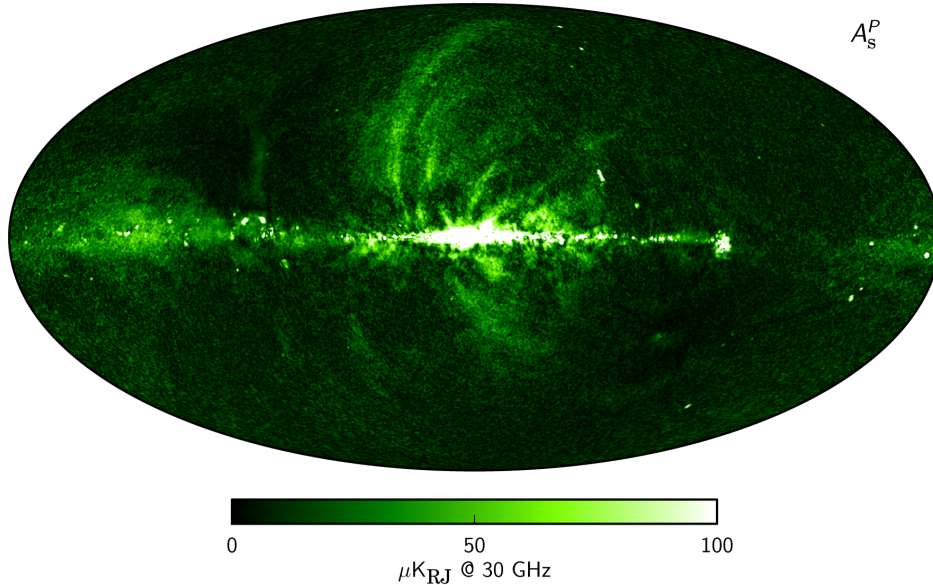
Figure 1.9: Planck synchrotron polarization amplitude map at 30 GHz, where $P = \sqrt{Q^2 + U^2}$ produced from Planck and WMAP data. Image from Ref. [17].

which deflects the photons from their straight line trajectories. While the lensing does not affect polarization direction or intensity, it does remap the position of temperature and polarization anisotropies. Gravitational lensing distorts E-modes into B-mode polarization which follows the frequency scaling of primordial B-modes (due to the photon source being the CMB), therefore a multi-frequency observation strategy will not remove these B-modes. The power spectra of these B-modes however follows the multipole scaling of CMB E-mode power spectra. This B-mode signal that is present at all frequencies must be delensed using the gravitational lensing potential. [20]

While we have a strong model for the lensing BB spectrum, there is still a variance contributed to the CMB BB power spectrum estimate. It is important to account for the lensing B-modes so that they are not mixed in with the B-mode signal of the CMB. A joint analysis by BICEP/Keck and South Pole Telescope [21] has shown that as we accumulate more data and the sensitivity of CMB measurements increases the lensing effects are more prevalent and delensing becomes necessary.

Chapter 2

BICEP/Keck Experiments

2.1   Program Overview

The BICEP/Keck experiment observes the CMB at low ℓ with the intent to detect polarization caused by PGWs, focusing on getting a measurement of tensor-to-scalar power ratio r. The BICEP/Keck telescopes are located in Antarctica at the Amundsen-Scott South Pole Station. Data has been taken here beginning in 2006 with BICEP1, continuing to BICEP2 in 2010, Keck Array in 2011, BICEP3 in 2015, and the most recent BICEP Array, which began taking data in 2020.

BICEP1 was the first telescope of the BICEP/Keck experiment, and took data from 2006 - 2008. The data analyzed by this telescope put a constraint of $r < 0.7$ [22]. BICEP2 was the first telescope to detect degree scale B modes in 2014, but it was determined after comparison with Planck data that these could not be disentangled from foregrounds [23]. Keck Array is a series of 5 telescopes of the similar design as BICEP2 on one mount and finished taking data in 2018. Analysis of Keck data combined with BICEP2, Planck and WMAP data produced a constraint of $r < 0.036$ [24]. BICEP3 began taking data in 2015 and is still taking data currently. This is a larger telescope than the detectors in Keck Array,

and takes data in the 95 GHz frequency band.

BICEP Array is the newest telescope, and is comprised of four BICEP3-scale receivers and is located where the Keck Array mount used to be. Currently there are two BICEP Array receivers deployed in the South Pole, with two spots in the BICEP Array mount containing Keck receivers for the time being. The BICEP Array receivers bring a new contribution to the experiment with detectors observing in previously unobserved (by BICEP/Keck) 30 and 40 GHz frequency bands in addition to adding more integrated sensitivity to the 95, 150, 220, and 270 GHz frequency bands [25].

2.2 Observing Site

The BICEP/Keck telescopes are stationed at the Amundsen-Scott South Pole Station which for many reasons is a near ideal location for ground-based CMB observations. Its location on the axis of Earth's rotation allows for the constant observation of the same area of sky, and gives the schedule of a one day cycle per year, allowing us to observe without interruption from the sun from February to November. There is little variation in the atmosphere and low precipitable water vapor at Pole to interfere with our observations.

The Amundsen-Scott South Pole Station is funded by the United States Antarctic Program which provides lodging, generators, travel options, and more for the staffing that stays there. The BICEP/Keck collaboration does not need to worry about these aspects which makes the South Pole a relatively simple place to run such an experiment. The telescopes are located in the Dark Sector, an area where there are no wireless signals located ∼1 km from the geographic South Pole. BICEP1, BICEP2, and BICEP3 have all been located in the Dark Sector Laboratory (DSL), and the Keck and BICEP Array telescopes have been located in the Martin A. Pomerantz Observatory (MAPO).

2.3 Instrument Design

While the current running telescopes are BICEP3 and BICEP Array, I will focus on describing the retired BICEP2/Keck Array receivers in this section as the simulations and data used in my analysis come from those experiments.

2.3.1 Instrument Overview

The BICEP2 telescope was built specifically to measure CMB polarization on 1 to 5 degree angular scales (or $\ell = 40 - 200$) [26] because CMB B-mode polarization is expected to peak within this range. The signal of B-modes compared to that of the total CMB polarization is such that the limiting factors of detection are systematics and sensitivity.

BICEP2 and Keck Array were built with on-axis refractive optics kept at a very low temperature to minimize internal loading, and a 26 cm aperture. The aperture was small enough to resolve degree-scale features in the sky while also preserving light-collecting power with a wide field-of-view. It had a detector design such that for each camera "pixel" there was a pair of beam-defining slot antennas coupled to photon-noise limited transition-edge sensor (TES) bolometers (kept at sub-kelvin temperatures, where each antenna had its own TES detector) which allowed for high sensitivity. Each member in an antenna/detector pair was labeled "A" and "B", where the the antenna pair is co-located on the focal plane such that each antenna is nominally observing the same area of sky. Physical detectors in reality are not sensitive to an infinitely fine point source, rather are simultaneously sensitive to a finite area about that source. The beam is the response of the detector to a point source, i.e. a point in the sky when looking at the CMB. There is some smoothing of the data that occurs from the convolution of the sky measurements by the beam, analogous to the point spread function for optical telescopes. Further description of the BICEP/Keck beam characterization is described in section 2.6.

The TES bolometers were produced by JPL Microdevices Laboratory as square silicon tiles containing an 8×8 array of dual-polarized spatial pixels, and supported multiplexed readout such that a total detector count of 500 bolometers at 150 GHz allowed greater sensitivity to B-modes than had been achieved with BICEP1. The focal plane contained 4 such tiles, totaling 500 optically coupled detectors and 12 "dark" (no antenna) TES detectors.

The antenna-coupled TES arrays were able to be densely packed with detector readout implementing multiplexing SQUID amplifiers [1] which reduced the amount of wires required, which therefore minimized the heat load on the focal plane.

Figure 2.1 shows a cross-sectional view of a BICEP2/Keck Array telescope insert. It consists of two tubes, with an optics tube on top and the camera tube below. In the optics tube, from top to bottom, are the objective lens, nylon filter, and eyepiece lens which are kept at 4 Kelvin. The superconducting components of the telescope are sensitive to ambient magnetic fields, such as those from nearby electrical equipment. These are shielded with a Nb magnetic shield kept at 350 mK. The focal plane assembly, the detector tiles pictured within the focal plane assembly, and the passive thermal filter are kept at 280 mK. Down in the camera tube are the flexible heat straps and the He sorption refrigerator, the fridge mounting bracket and the camera plate which are kept at or below 4 K.

The bottom plate of the insert was directly connected to a helium bath (the Keck Array used pulse tube coolers rather than the helium bath) which kept the whole insert cooled to 4 K, and allowed the sorption fridge to condense liquid helium. The detector tiles were mounted to the focal plane, built around a gold-plated, oxygen-free high thermal conductivity copper detector plate which had four square windows allowing radiation to reach the detectors (shown in figure 2.2).

The BICEP2 lenses were made of high density polyethylene with a diameter of   30

---

[1]Superconducting Quantum Interference Device (SQUID) multiplexers allow for building arrays of thousands of low-temperature bolometers and microcalorimeters with a reasonable amount of readout channels [27]
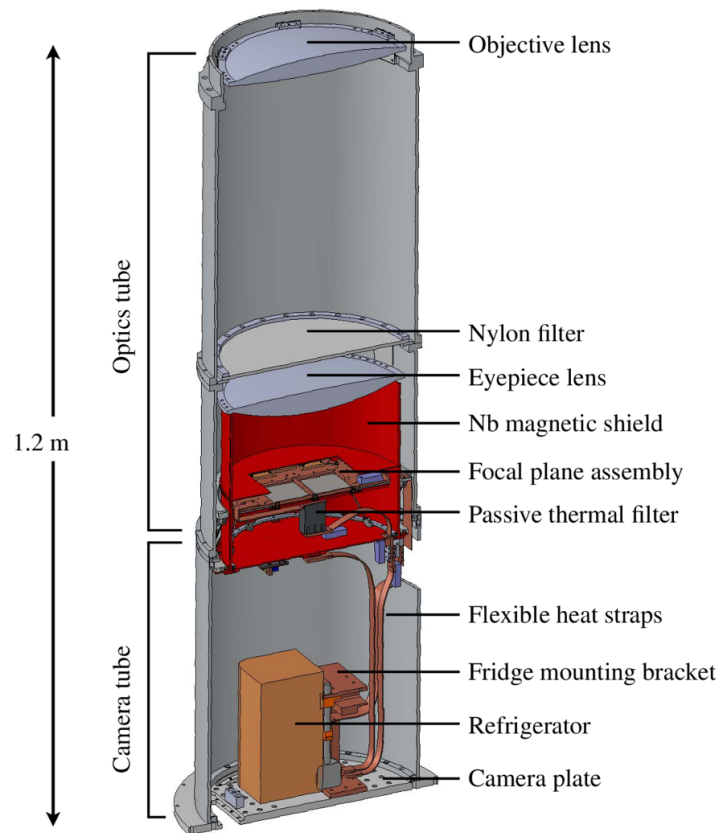
Figure 2.1: BICEP2/Keck Array receiver cross-section [26]. The optics tube is a rigid structure containing the focal plane assembly, filters, and lenses. The camera tube encloses the refrigerator and and cryogenic readout electronics.
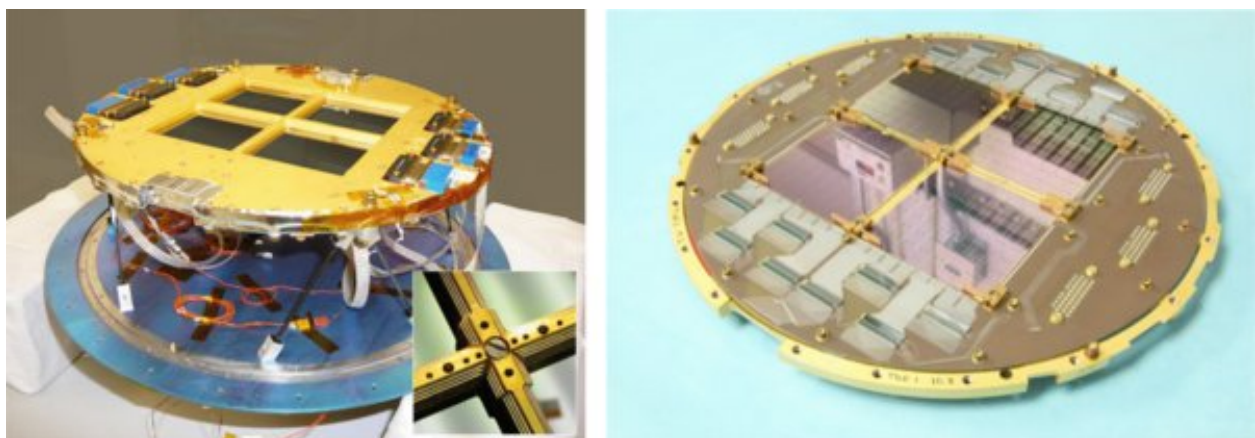
Figure 2.2: BICEP2 focal plane assembly on the carbon-fiber truss structure and a 350 mK Nb plate. On the left is the unsheilded assembly, and on the right is the underside of the focal plane Cu plate. [26]
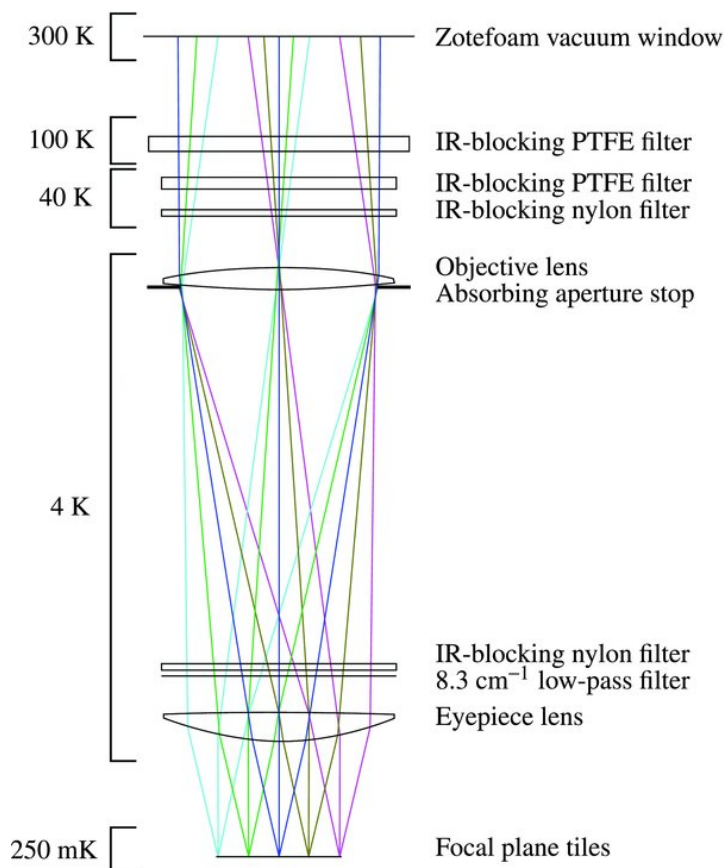
Figure 2.3: BICEP2 and Keck Array Optical Design [26]. Besides the vacuum window, all components have an antireflection coating for minimal reflection at 150 GHz. The optics below the nylon filter were kept cooled at 4K for low, stable optical loading.

cm. The optical design allowed for an effective focal length of 587 mm and a 550 mm lens separation. Simulations of this design predict well-matched beams for two idealized orthogonally polarized detectors on the focal plane. Figure 2.3 shows the BICEP2 optical design.

2.4 Observing Strategy

All generations of the BICEP/Keck telescopes follow a similar observation strategy. The same patch of sky is always observed, an area referred to as the "Southern Hole" outside of the galactic plane, which has low foreground emissions and accounts for 1% of the sky.
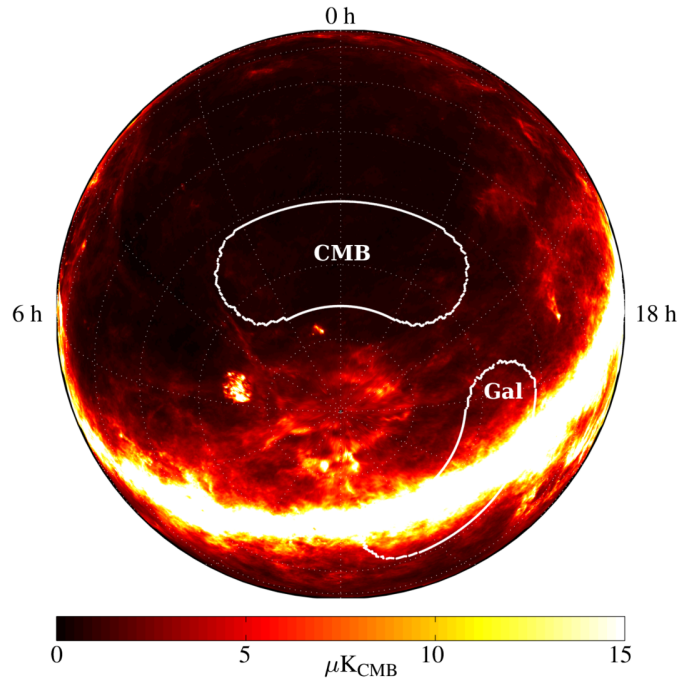
Figure 2.4: The patch of sky observed by BICEP/Keck referred to as the Southern Hole, with background color begin a map of dust brightness at 150 GHz. Image from Ref. [26].

The telescopes scan in RA, step up in Dec, and repeat the azimuth scan repeatedly. For BICEP2 and Keck generation telescopes, the observations were centered at RA = 0h and Dec = $-57.5°$ Some understanding of terminology is required for the remainder of this chapter.

- **Elnod:** An "elevation nod" is a 1.28° peak-to-peak triangle wave scan performed about the central elevation of a scanset in fixed azimuth. Each scanset has a leading and trailing symmetric up-down-up and down-up-down elnod. Since the atmosphere is unpolarized this motion is done as a relative gain calibration.

- **Halfscan:** The smallest movement of the telescope during a CMB observation, a halfscan is a smooth 64.4° sweep of the telescope in azimuth at constant elevation. The central 50° of this scan is at a constant speed of 2.8°/s, with a smooth deceleration/acceleration at the turnaround in preparation for the next halfscan.

- **Scanset:** A scanset is comprised of ∼50 pairs of left and right-going halfscans at fixed elevation and azimuth center, allowing the sky to drift 12.5° per scanset. Before

and after each scanset is each variation of elnod and a partial load curve [2] for data calibration. The elevation and azimuth center positions are updated every other scanset to account for the turning sky and to provide more sky coverage.

- **Phase:** A phase is a collection of 6-10 scansets. Each phase is labeled by a letter of the alphabet from A through I, and covers different ranges of azimuth and elevation.

2.5 Data Reduction

Time Ordered Data

The raw time-ordered data during a CMB observation is directly sent from Pole and stored to disk. This is the lowest level of reduced data and is monitored by graduate students and discussed on a weekly basis within the collaboration. This is where small details are initially observed, and is important for the experiment to run smoothly. Constant monitoring of the incoming data is where we check for anything that could become a problem, and is very important in a collaboration such as BICEP/Keck with an experiment so remotely located with limited communication. Figure 2.5 shows an example of what the timestream data looks like. The low-level processing steps applied to this data are described below.

Deconvolution

Because the data has to be sent from Pole to North America, in the interest of efficiency the data is down-sampled. Therefore there is a series of down-sampling by MCE [3], and filtering by GCP [4] that the data goes through between readout at the bolometers and being written

---

[2] A partial load curve (PLC) is a voltage sweep of the detector bias, in which we begin at high bias and and end at the target bias, giving us an IV curve.

[3] Multi-Channel Electronics (MCE) is a readout and biasing system developed by the University of British Columbia [28]

[4] Generic Control Program (GCP) is a software control system that reads out and stores the data from the telescopes
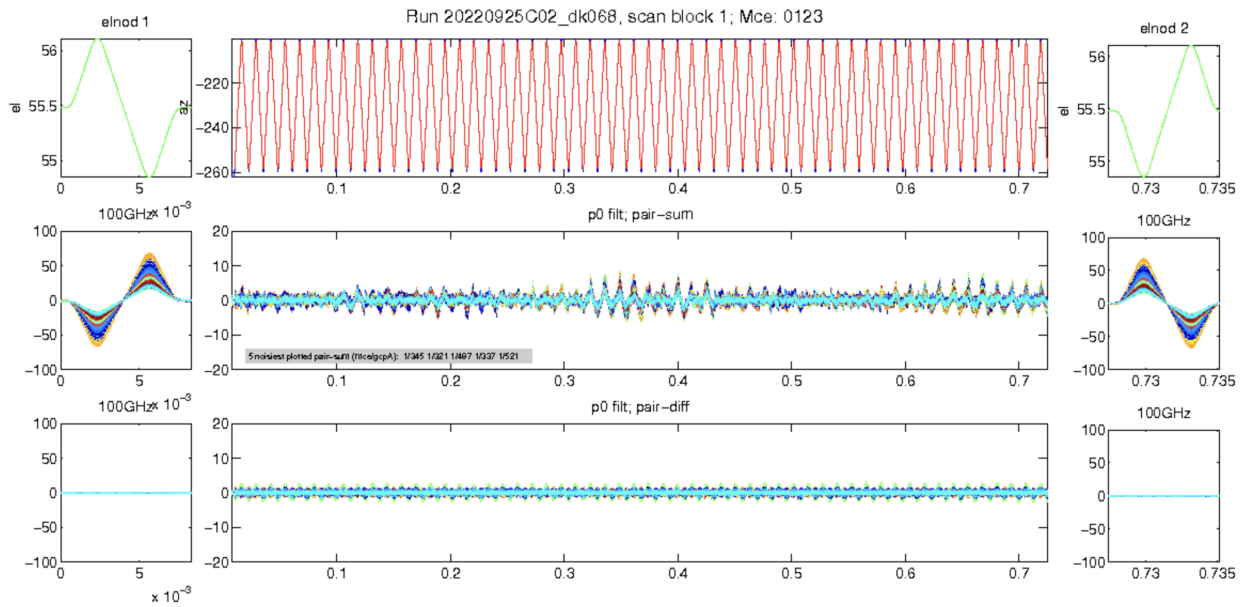
Figure 2.5: BICEP3 data for one scanset. On the x-axis is time in fractions of hours, such that the full 50 minute scanset is shown. The top center panel shows the motion of the mount. The center panel shows the pair-sum data, which measures the total intensity of the incoming radiation, and the bottom center shows the pair-diff data, which shows the polarized component. On either side of each of the center panels are plots for the leading and trailing elnods. The data shown is after round2 cuts have been performed and is in arbitrary feedback units.

to disk. The first step of looking at the data is then to deconvolve the transfer functions of these downsampling and filtering steps.

De-glitching

The next step in data reduction is de-glitching. Our data is often subject to spikes and jumps due to the high gain of the detectors and readout system. Spikes can occur in the TOD timestreams from cosmic ray hits and are removed along with the data one second before and one second after the event which is set to NaN. Discontinuous jumps in the data can occur from a flux-jump in a SQUID, causing it to jump from one lock-point to another. This data is removed, and the data right before the glitch is matched to the data immediately after.

Relative Gain Calibration

As mentioned in section 2.4, there are elnods performed at the beginning and end of each scanset. This motion injects a signal into the detector timestreams proportional to the atmospheric opacity gradient [26]. The data from each detector is read out in arbitrary units (FBU) and as such each detector does not necessarily have the same gain. The gain for each channel can be roughly equalized with the derivation of a relative gain (relgain) coefficient from SQUID feedback units per air mass, which comes from regressing the elnod against the air mass profile through which it was pointing. Dividing each scanset's timestream by its own relgain coefficient and multiplying by the median over all good detectors does the equalizing, as well as canceling out a significant amount of atmospheric fluctuations when taking the difference of the detector pairs which is very important to the construction of maps (as will be discussed in section 2.6). The relgain coefficients are obtained in units of FBU per airmass, and later in section 2.6 the FBU units will be converted to units of μK_{CMB}.

Data cuts

Before the data is able to be used to create maps, a series of selections and cuts is done to remove bad quality data due to weather and atypical instrument behavior. This data selection process is divided into two steps, Round 1 cuts and Round 2 cuts. The Round 1 cuts are performed on halfscans, and Round 2 cuts are done at each scanset.

Timestream Filtering and Ground Subtraction

Timestream filtering is when each halfscan is fit with a third ordered polynomial which is then subtracted to account for changing weather conditions. An average of each hour of data is taken in azimuth as a ground template and is removed from each halfscan.

2.6   Map Making

From TOD to Map

Here we describe the process of converting our weighted time ordered data to two-dimensional binned pixel maps in the celestial coordinates RA and Dec [29]. The TOD consists of the sum and difference of the signals observed by the orthogonal detector pairs A and B as described in section 2.3.1. Each detector is sensitive to both unpolarized T and polarized Q and U, with observed timestream of detector A being:

$$\tau_t^A = T_t + cos(2\Psi_t^A)Q_t + sin(2\Psi_t^A)U_t \tag{2.1}$$

where T_t, Q_t and U_t are the temperature and Stokes parameters for timestream sample t, and there are n_t time ordered measurements of the sky within a timestream. Ψ_t^A is the angle that the detector A makes with the Q,U axis in the sky. The sum and pair difference then

is taken with respect to detector B:

$$s_t = \frac{1}{2}(\tau_A + \tau_B) = T_t + \alpha_t^+ Q_t + \beta_t^+ U_t$$
$$d_t = \frac{1}{2}(\tau_A - \tau_B) = \alpha_t^- Q_t + \beta_t^- U_t$$

$$(2.2)$$

where α_t and β_t are defined with $\Psi_t^A(\Psi_t^B)$ being the angle the A(B) detector makes with the Q,U axis on the sky:

$$\alpha_t^{\pm} = \frac{1}{2} \left[cos(2\Psi_t^A) \pm cos(2\Psi_t^B) \right]$$
$$\beta_t^{\pm} = \frac{1}{2} \left[sin(2\Psi_t^A) \pm sin(2\Psi_t^B) \right]$$

$$(2.3)$$

If we make the assumption that the detector pairs A/B are co-pointed and perfectly orthogonal to one another, we can rewrite the pair sum and pair differences in equation (2.2) as

$$s_t = A_{tj} T_j$$
$$d_t = \cos(2\Psi_t^A) A_{tj} Q_j + \sin(2\Psi_t^A) A_{tj} U_j$$

$$(2.4)$$

where A_{tj} is the transformation from the input pixel map (with pixel index j) to the timestream sample (with index t), and α_t^+ and β_t^+ cancel in the pair-sum signal due to the orthogonality. We see now that the temperature map simply comes from the pair-sum while the pair-diff gives a linear combination of the Q and U maps.

Taking into account the fact that we take many observations over a single pixel, we sum over the N samples during which process we weight each detector pair i (w_i) with the inverse timestream data during the scanset. We can then in matrix form compute T, Q and U maps by the following:

$$\sum_i w_i = \begin{bmatrix} s_i \\ \alpha_i^- d_i \\ \beta_i^- d_i \end{bmatrix} = \sum_i w_i \begin{bmatrix} N & 0 & 0 \\ 0 & (\alpha_i^-)^2 & \alpha_i^- \beta_i^- \\ 0 & \alpha_i^- \beta_i^- & (\beta_i^-)^2 \end{bmatrix} \begin{bmatrix} T \\ Q \\ U \end{bmatrix} \tag{2.5}$$

And then obtain the T, Q and U maps through the matrix conversion:

$$\begin{bmatrix} T \\ Q \\ U \end{bmatrix} = \left(\sum_i w_i \begin{bmatrix} N & 0 & 0 \\ 0 & (\alpha_i^-)^2 & \alpha_i^- \beta_i^- \\ 0 & \alpha_i^- \beta_i^- & (\beta_i^-)^2 \end{bmatrix} \right)^{-1} \left(\sum_i w_i \begin{bmatrix} s_i \\ \alpha_i^- d_i \\ \beta_i^- d_i \end{bmatrix} \right) \tag{2.6}$$

Multiple detector angles are required for this matrix to be invertible, which is taken care of physically by rotating the boresight of the telescope through multiple angles.

Deprojection

Ideally, the detector pairs A and B are orthogonal to one another, but in reality this is not always the case. Small mismatches in detector pairs can cause temperature to polarization leakage, and because the intensity of temperature is so much brighter than that of polarization this can overwhelm the polarization maps and cause false B-mode signal. The beam of each detector in a pair can be modeled as an elliptical Gaussian, and the difference between these Gaussian curves can be Taylor expanded which to second order can be used in different deprojection techniques to filter out this $T \rightarrow P$ leakage.

Three types of leakage occur and are described in more detail in the BICEP2 systematics paper [30]:

1. **Monopole:** If the detectors Gaussian heights or widths are mismatched (difference in gain between detectors), the leakage caused by these mismatches is monopole symmetric, meaning the differences are invariant under rotation. The false polarization signal

from $T \to P$ leakage changes sign once the detector pairs are rotated $90°$ due to the spin-2 nature of the polarization fields. In other words, under $90°$ rotation leakage of temperature to polarization switches sign, $Q \to -Q$ and $U \to -U$. Because of this, monopole leakage can be canceled out through averaging over signal collected at deck rotations of $90°$.

2. **Dipole:** Mismatched beam centroids among detector pairs causes dipole symmetric $T \to P$ leakage. Similarly to monopole leakage, dipole leakage changes sign under a $180°$ deck rotation, and again can be canceled out through multiple observations taken through $180°$ rotations.

3. **Quadrupole:** A mismatch in the relative angle of ellipses can cause the most problematic leakage, quadrupole symmetric leakage. This type of leakage looks like real CMB polarization which is also quadrupole symmetric, and no sequence of boresight rotation can cancel out this type of leakage. Because of this, the real beam mismatch must be accurately simulated to predict the effects of this ellipticity mismatch in the data.

While observing under the different deck rotations cancels out a good amount of the monopole and dipole symmetric leakage, not all is removed because there is some unequal detector and deck coverage. Deprojection is performed to filter out this leftover leakage.

Each of the beam modes can be described with an elliptical Gaussian of the form:

$$B(x) = \frac{g}{\Omega} \exp \left[-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^- (\mathbf{x} - \mu) \right] \tag{2.7}$$

where $\mathbf{x}$ is a two dimensional position vector with relation to the focal plane. This coordinate system is defined for each detector pair such that two orthogonal great circles intersect at the centroid of the detector beams, and rotates with the focal plane under deck angle rotations. μ is the position of the peak, g is the gain and Ω a normalization factor, and Σ is a covariance

matrix defined:

$$\Sigma = \sigma^2 \begin{bmatrix} 1+p & c \\ c & 1-p \end{bmatrix} \tag{2.8}$$

where σ is the beam width and c and p are ellipticity parameters describing "plus" and "cross" ellipticity. These describe the allignment of the ellipse major axis with the x axis and $45°$ from the y axis of the focal plane. $+(-)p$ ellipticity has the major axis pointing along the $x(y)$ axis, and $\pm c$ ellipticity has the major axis at $\pm 45°$ to the x axis.

If we have two ideal detectors A and B then we can model them as circular Gaussian beams ($p = c = 0$) with width σ that are pointed in the same direction ($\mu = 0$). If we perturb (δ) one of the parameters (k) and take the differences between the two beams, we can then Taylor expand about the parameter of interest to first order and obtain six differential beam modes $B_{\delta k}$ (where $\delta k = k_A - k_B$) [30]:

$$\begin{aligned} B_{\delta g} &= \delta g B(x,y) & B_{\delta \sigma} &= \sigma \delta \sigma (\nabla_x^2 + \nabla_y^2) B(x,y) \\ B_{\delta x} &= \delta x \nabla_x B(x,y) & B_{\delta y} &= \delta y \nabla_y B(x,y) \\ B_{\delta p} &= \frac{\sigma^2}{2} \delta p (\nabla_x^2 - \nabla_y^2) B(x,y) & B_{\delta c} &= \frac{\sigma^2}{2} \delta c (2\nabla_x \nabla_y) B(x,y) \end{aligned} \tag{2.9}$$

The three types of $T \rightarrow P$ leakage can be described by these six parameters and characterized with their symmetries: monopole ($B_{\delta g}$ and $B_{\delta \sigma}$), dipole ($B_{\delta x}$ and $B_{\delta y}$), and quadrupole ($B_{\delta p}$ and $B_{\delta c}$). Figure 2.6 shows a visualization of $T \rightarrow P$ leakage coming from these differential beam modes.

The convolution theorem allows us to move derivatives of the beam modes to those of a temperature map to get a deprojection template which is then used to perform deprojection:

$$\left(\frac{\partial}{\partial \eta} B\right) * T = \frac{\partial}{\partial \eta} (B * T) = B * \left(\frac{\partial}{\partial \eta} T\right) \tag{2.10}$$
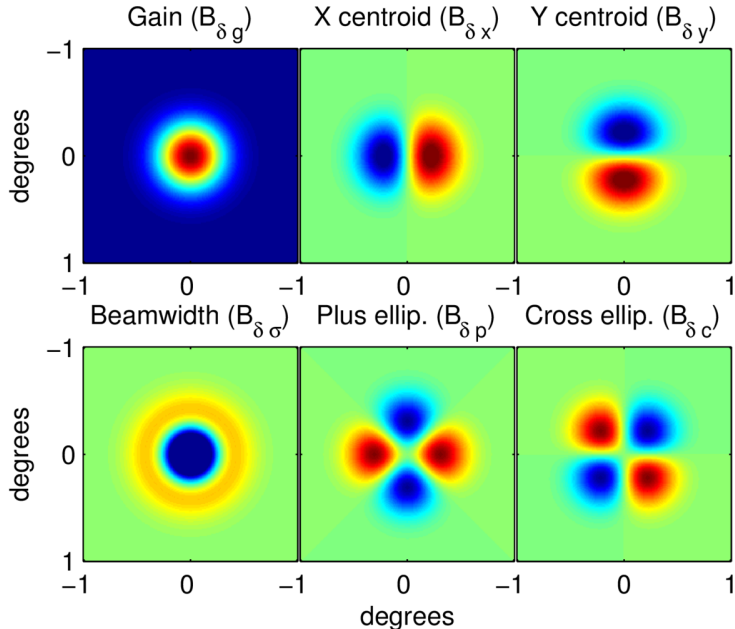
Figure 2.6: Difference in elliptical Gaussian beams. The gain and beamwidth panels show monopole leakage, centroid X and Y panels show dipole leakage, and the Plus and Cross ellipses plot shows quadrupole leakage. [30]

We can collect these template maps for each detector on a per-scanset basis. We build simulated timestreams by sampling the templates along the detector pointing location. The timestreams are then filted and binned into pairmaps exactly as is done for the real data. We obtain the deprojection coefficients by simultaneously regressing all of the six templates against the real pair-difference time streams.

For the monopole and dipole leakage, we can simply apply the deprojection by scaling the template by the necessary coefficient and subtract from the data. This causes some loss in data which can be precisely accounted for in simulation. In the case of quadrupole leakage, we must use direct measurements of the beam mismatch to scale the templates which we subtract from the data.

Absolute Calibration

Until now, the data has been stored in arbitrary units and has undergone a relative gain calibration. At this point the data is ready to undergo an absolute calibration which will convert the maps from their arbitrary units to more useful CMB temperature units, μK_{CMB}. An absolute calibration factor, or abscal, is computed by first cross-calibrating our observed temperature map with a reobserved Planck reference temperature map, which has undergone a BICEP/Keck beam smoothing and filtering. We then compute another cross spectrum between this same reference map and another calibration Planck map generally chosen to be close in frequency to the observed map that we are calibrating. The ratio between these two cross spectra gives us a conversion factor of $\mathrm{FBU}/\mu K_{CMB}$. This factor g_b is calculated for each ℓ bin (further described in section 2.7):

$$g_b = \frac{\langle m_{ref} \times m_{cal}\rangle_b}{\langle m_{ref} \times m_{uncal}\rangle_b} \tag{2.11}$$

The reobserved Planck reference map is denoted as m_{ref} and the calibration Planck map close in frequency to our uncalibrated (m_{uncal}) is denoted m_{cal}.

2.7 Power Spectra

Once we have our data binned into calibrated maps of the sky we are ready to calculate angular power spectra, but first need to perform a bit of manipulation. Because as described in section 2.4 we observe a 1% patch of sky, we can use flat sky approximation and convert our maps into Fourier space. Before doing this, we need to weight the pixels by their inverse noise variance in order to obtain a clean Fourier map.

The first step is calculating the variance of our pixel map and then taking the inverse. Doing so down-weights the edges of the maps as these are more noisy due to having less

observation time than the center of the map. This also provides a smooth transition to zero at the edge of the map allowing for a well-behaved transform. This inverse variance map is then multiplied by our pixel map to create an apodized map, which then undergoes a 2D Fourier transform. We now have temperature and polarization maps in Fourier space denoted $\hat{T}(\ell_x, \ell_y)$, $\hat{Q}(\ell_x, \ell_y)$, and $\hat{U}(\ell_x, \ell_y)$, where ℓ_x and ℓ_y are the coordinates in Fourier space analagous to the wavenumbers k_x and k_y.

We recall that our $\hat{Q}$ and $\hat{U}$ maps are coordinate dependent, and can convert to our $\hat{E}$ and $\hat{B}$ maps as described in section 1.2.2. These maps are most useful to us in radial coordinates (ℓ, ϕ) using $\ell = \sqrt{\ell_x^2 + \ell_y^2}$ and $\phi = tan^{-1}(\frac{\ell_x}{\ell_y})$, and this rotation from $\hat{Q}$ and $\hat{U}$ to $\hat{E}$ and $\hat{B}$ is shown in equation 2.12.

$$\hat{E}(\ell) = \hat{Q}(\ell)cos(2\phi) + \hat{U}(\ell)sin(2\phi)$$
$$\hat{B}(\ell) = \hat{Q}(\ell)sin(2\phi) - \hat{U}(\ell)cos(2\phi)$$

(2.12)

The Fourier apodized map is multiplied by its complex conjugate, giving us a two dimensional angular power spectrum, an example of which is shown in figure 2.7. The units of the x and y axes of this map are rescaled from 1/deg to approximately ℓ, and the z-axis is in units of μK^2. On the left hand side of each plot are the square of the imaginary parts, and on the right are the square of real parts. There are 17 annuli with radii ranging from $\ell \sim 10$ to $\ell \sim 600$ which is how we bin our ℓ values, and each annulus has a width of $\Delta \ell = 35$. The "science" bins which are studied in the BICEP/Keck analysese are the second through tenth bins. The two left panels are the $\hat{Q}$ and $\hat{U}$ speclin maps, and the right panels are the $\hat{E}$ and $\hat{B}$ speclin maps. The points in the $\hat{E}$ map are much brighter than the $\hat{B}$ map as is expected, and these points correspond to the x and y axes of the $\hat{Q}$ map and 45° rotated in the $\hat{U}$ map.

We calculate an estimate of the power spectrum for each bin by averaging over the points in each annulus, and the binned power spectra, or pseudo-D_ℓ are referred to as bandpowers.
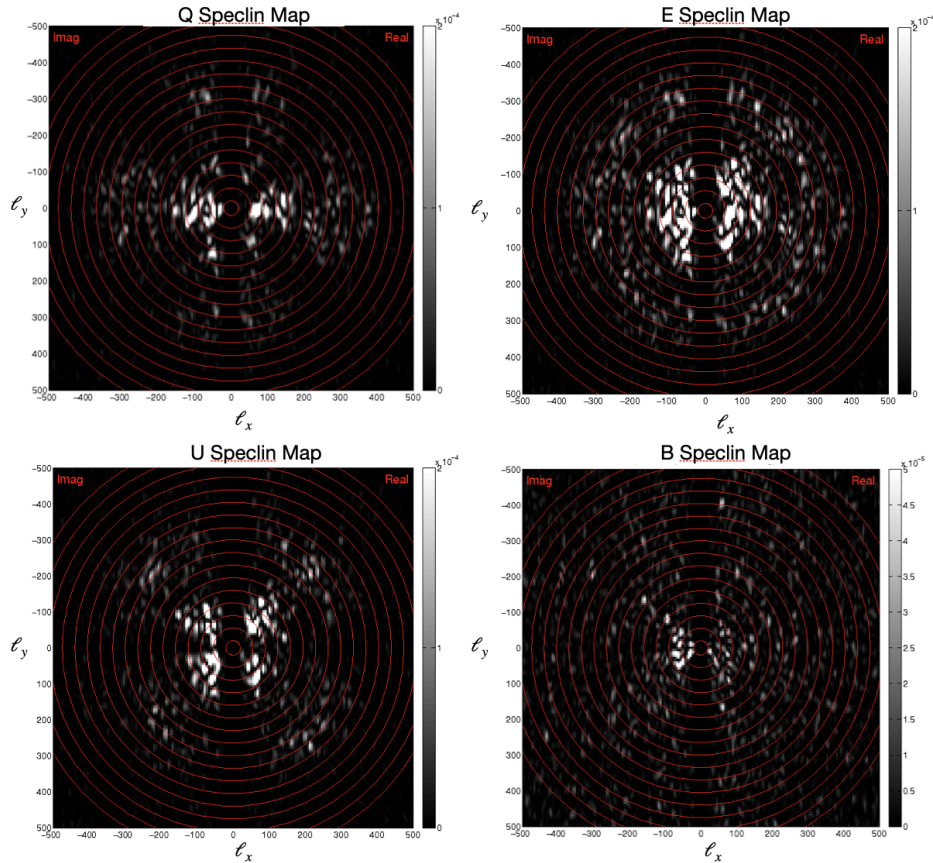
Figure 2.7: Square of Fourier maps from BK15, where the left panels show the Stokes parameters maps, and on the right are the E and B-mode Fourier maps. The left side of each plot shows the imaginary and on the right are the real values. The E-modes are dominant, and can be seen on the x and y-axes of the Q map, and on the 45° diagonals of the U map. The gap that can be seen close to $\ell_x = 0$ is due to the third order poly subtraction filter and ground subtraction. The red annuli represent the bins for which we calculate the power spectra.

The filtering from the data processing and the beam as well as the map apodization suppresses and mixes the signal between modes. We account for this by calculating bandpower window functions for each ℓ bin for each frequency band power spectrum, which relate the theoretical input power spectrum to our bandpowers:

$$\langle D_{b,\ell} \rangle = \sum_{\ell} w_{b,\ell} D_{\ell} \tag{2.13}$$

where $\langle D_{b,\ell} \rangle$ is the expectation value for our bandpowers, D_{ℓ} is the theoretical power spectrum, and $w_{b,\ell}$ are the BPWF for each bin. Our process of apodization mixes signal across angular scales, and shows up in the shape of the BPWF. Figure 2.8 shows the 150 GHz BPWF for our 9 science bins, and we see this mixing as the contribution to the lower ℓ bins from the higher ℓ BPWF. To calculate the BPWF, 100 map realizations are constructed for each multipole with unit power and then reobserved. The integral of these BPWF gives us a suppression factor for each bin which tells us how much power was lost in the data to map process. We can obtain a BPWF and suppression factor for both auto power spectra and the cross power spectra between frequency bands.

2.8   WMAP and Planck Maps

The BICEP/Keck collaboration ground based data has been studied in conjunction with the full sky maps made by the outer-space based experiments WMAP and Planck. WMAP was a 9 year mission that was launched June 30 of 2001 and ended August 2010 [31] with the intent to replicate and confirm the COBE sky maps, as well as measure the sky beyond what COBE had done. The WMAP mission observed the sky through 5 different frequency bands surrounding the region within which the CMB to foreground signal to noise is maximized. The two frequency bands from the WMAP data set used for the BICEP/Keck data analysis are the 20 and 23 GHz frequency bands.
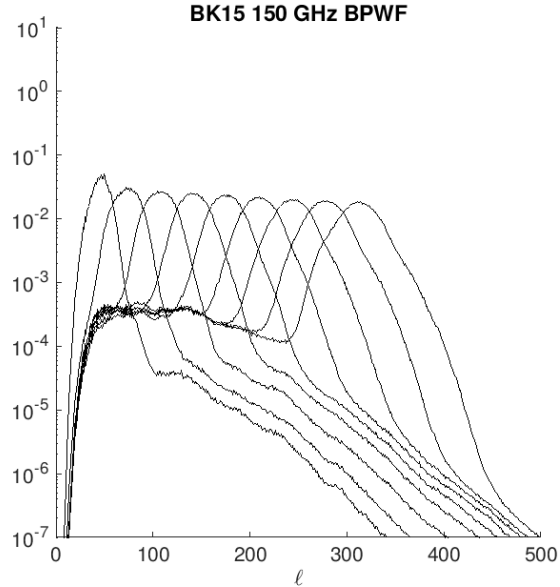
Figure 2.8: The BK15 150 GHz bandpower window functions for the 9 science bins. We can see each BWPF centered about each bin, however especially at high ℓ the BPWF have some contribution at lower ℓ. We note that the scale of the y-axis is logarithmic.

The majority of the external maps utilized in the BICEP/Keck data analysis come from the Planck satellite, which observed continuously from August of 2009 to October of 2013 [15]. Planck was equipped with detectors sensitive to 9 frequency bands between 25 GHz and 1000 GHz. The Low Frequency Instrument (LFI) had detectors which covered frequency bands centered at 30, 44, and 70 GHz while the High Frequency Instrument (HFI) detectors covered the frequency bands centered at 100, 143, 217, 353, 545, and 857 GHz. All frequency maps by Planck with exception of the 545 and 857 GHz, which did not have polarization sensitivity, were analyzed with the WMAP and BICEP/Keck maps.

Because the WMAP and Planck maps cover the entirety of the sky while the BICEP2 and Keck Array maps cover only 1% of the sky, we need to get the external maps in a format to be usefully integrated into the BICEP/Keck data set. Not only is the sky coverage vastly different between the experiments, BICEP/Keck also has beams and filtering specific to the experiment which also affect how we can compare our maps to the external ones. To account for these differences, we need to "reobserve" the external maps according to the

BICEP/Keck framework.

Because the BK15 150 GHz maps are the least noisy, we reobserve the external maps with the BK15 150 GHz observation matrix. The observation matrix includes all of the filtering operations so that it can transform to the true sky to the sky as observed by BICEP/Keck through the 150 GHz frequency. The steps required to do this are to first deconvolve the WMAP(or Planck) beam and reconvolve with the BK15 150 GHz beam. We then rotate from galactic to celestial coordinates and truncate the full sky external maps to the pixels which are compatible with the observing matrix, and then finally multiply each map by the observing matrix.

In summary, there are a total of 12 frequency maps which are contained in the BK15 dataset:

- 3 Planck LFI maps: 30, 44, and 70 GHz

- 4 Planck HFI maps: 100, 143, 217, and 353 GHz

- 2 WMAP maps: 20 and 33 GHz

- 3 BICEP/Keck maps: 95, 150, and 220 GHz

From these maps we can obtain a 12×12 covariance matrix containing all of the auto and cross power spectra between the maps where the auto power spectra are located on the diagonal. The order in which the maps are multiplied by each other to calculate the cross spectra does not matter, so this is a symmetric covariance matrix, giving us a total of 78 unique power spectra. The covariance matrix is integral to the analysis detailed in the following chapters, in which I will detail the an internal linear combination component separation method that takes a weighted average of all of the frequency maps listed above, and the covariance matrix is used to calculate the weights that are applied to the maps.

Chapter 3

Internal Linear Combination Component Separation Method

3.1 Motivation

The advancement of technologies specializing in observing the CMB has vastly improved measurements of the temperature and polarization fluctuations to such a degree that foregrounds have become the greatest source of contamination in CMB observations. Due to this, developments in data analysis methods which aim to identify and separate foreground contamination from CMB signal have become of utmost importance.

One of the most common class of algorithms for foreground disentanglement is called Internal Linear Combination, a component separation method with variations used by other experiments such as Planck [32], which uses a Scale-discretized ILC variation and WMAP [33], for which a modified internal linear component algorithm (MILCA) variation was developed that includes extra steps such as correcting for intrinsic noise bias. The ILC method averages over measurements of the sky taken at different frequencies, with the aim to jointly minimize noise and foregrounds while recovering the CMB signal. It is different than other

methods of separating foregrounds which use models of dust composition because it makes very little assumptions about the contents of sky maps beyond the fact that there exists a mixture of sources. The following sections describe the specific ILC method used in this analysis to recover CMB signal using simulations from the BICEP/Keck 2015 data release [34].

3.2 Internal Linear Combination

The Internal Linear Combination (ILC) method is a component separation method which averages over measurements taken at different frequencies. Component separation indicates a class of methods that aim to recover a specific signal from a mixture of sources. The purpose of the ILC is to reconstruct the signal of interest, the CMB in our case, by minimizing the total error variance between the estimated and true CMB signal [35]:

$$\chi^2 = \sum_{\vec{\ell}} |\hat{s}(p) - s(p)|^2 \tag{3.1}$$

Where $\hat{s}$ is the estimated signal component, s is the signal, and p is the space of interest for which these signal maps exist. For this analysis we work in Fourier space, with $\vec{\ell}$ as the position (ℓ_x, ℓ_y) in the Fourier plane. The ILC assumes that the CMB signal is the same for all frequencies up to a scaling factor, which in our units (μK_{CMB}) is 1.

In order to ensure that the CMB temperature is equivalent for all observed frequencies, we employ the thermodynamic temeprature unit T_{CMB} which is proportional to the brightness temperature T_B through the relation:

$$T_B = \frac{x}{\exp x - 1} T_{CMB}$$
$$x \equiv \frac{h\nu}{kT} \tag{3.2}$$

where T_B is the brightness temperature given by

$$T_B = \frac{c^2}{2k\nu^2} B(\nu) \tag{3.3}$$

and brightness is

$$B(\nu) = \frac{2h\nu^3}{c^2} \left[\exp\left(\frac{h\nu}{kT}\right) - 1 \right]^{-1} \tag{3.4}$$

and T is the measured temperature of the CMB.

Because the foregrounds are optically thin, we may model observation y with the following linear equation:

$$y_i(\vec{\ell}) = s(\vec{\ell}) + f_i(\vec{\ell}) + n_i(\vec{\ell}) \tag{3.5}$$

where y_i is the observed data in Fourier space coordinates, s is the true CMB signal, and f_i and n_i are the foreground signal and noise for each observed frequency i. The only assumptions made by the ILC are that the CMB signal is the same at all frequencies, and we can model our data as a sum of the different components, where foreground signal and noise do not correlate with the CMB signal.

The ILC takes a weighted combination of the measurements from each frequency to form one overall measurement, assigning a specific weight $w_{i,\ell}$ ($\ell = \sqrt{\ell_x^2 + \ell_y^2}$ being the multipole number) to each frequency band and multipole. This gives us the estimated signal $\hat{s}$ of the following form:

$$\hat{s}(\vec{\ell}) = \sum_i w_{i,\ell} \cdot y_i(\vec{\ell}) \tag{3.6}$$

We use the simplest form of the ILC, in which the weights minimize $\hat{\sigma}_{CMB}^2$, the variance

of the estimated signal $\hat{s}$. In order to preserve the true signal s, the weights must satisfy the condition:

$$\sum_i w_{i,\ell} = 1 \tag{3.7}$$

Inserting equation (3.5) for y_i, equation (3.6) becomes:

$$\hat{s}(\vec{\ell}) = s(\vec{\ell}) + \sum_i w_{i,\ell} \cdot f_i(\vec{\ell}) + \sum_i w_{i,\ell} \cdot n_i(\vec{\ell}) \tag{3.8}$$

With the variance:

$$(\hat{\sigma}^2_{CMB})_\ell = w_\ell^\dagger \mathbf{R}_\ell w_\ell \tag{3.9}$$

$\mathbf{R}_\ell$ is the covariance matrix of the frequency maps at scale ℓ, with the diagonal elements containing the auto power spectra of the frequency maps and the off-diagonals containing the cross power spectra between the maps. w_ℓ is a vector containing the weights w_i for all frequencies i at multipole ℓ.

Minimizing $\hat{\sigma}^2_{CMB}$ (under the constraint in equation (3.7)) also minimizes χ^2 in equation (3.1). The minimum of $\hat{\sigma}^2_{CMB}$ can be found by using Lagrange multiplication on the following system of equations:

$$\forall i \quad \frac{\partial}{\partial w_{i,\ell}}[(\hat{\sigma}^2_{CMB})_\ell + \lambda(1 - \sum_i w_{i,\ell})] = 0$$
$$\sum_i w_{i,\ell} = 1 \tag{3.10}$$

Which gives us the following equation for weight at each frequency i and scale ℓ:

$$w_{i,\ell} = \frac{\sum_j [\mathbf{R}_\ell^{-1}]_{i,j}}{\sum_{i,j} [\mathbf{R}_\ell^{-1}]_{i,j}} \tag{3.11}$$

We note that the derivations performed above are under the assumption that the noise and foreground maps have been calibrated with respect to the CMB component. Were this not the case, we could still perform the ILC process with a slight modification to equation (3.5):

$$y_i(\hat{\ell}) = a_i s(\hat{\ell}) + f_i(\hat{\ell}) + n_i(\hat{\ell}) \tag{3.12}$$

where a_i are the recalibration coefficients to the CMB signal. In this case then the weights equation would become:

$$\mathbf{w} = \frac{\mathbf{A}^T \mathbf{R}^{-1}}{\mathbf{A}^T \mathbf{R}^{-1} \mathbf{A}} \tag{3.13}$$

where $\mathbf{A}$ is the vector containing the a_i recalibration coefficients. Other variations of the ILC include breaking down the sky maps and performing ILC separately on each of the regions. The final ILC map is the addition of the independentaly obtained ILC maps.

3.3   ILC Adaptations for BICEP/Keck Pipeline

There are some details that must be taken into account when implementing the ILC with maps coming from the BICEP/Keck pipeline, making it more complicated than described above.

Although the ILC assumes that the CMB signal is the same for each observational frequency, this is not true of the maps that we have. Our maps have encoded filtering and beaming which is frequency dependent. We do not currently have the tools to make the

CMB signal the same at the map level, however we can use suppression factors (described in 2.7) for binned power spectra to do this at the power spectrum level. As such, the power spectra must be suppression factor corrected both in the input map covariance matrix $\mathbf{R}_\ell$ and in the final ILC power spectrum.

The weights we calculate depend on discrete ℓ values, so it is important to decide on a binning scheme. We use one with 17 ℓ bins due to the ready availability of suppression factors in the BICEP/Keck pipeline for this binning scheme. While the weights are calculated per bin, the maps that these weights are multiplied by are almost continuous in ℓ (with the caveat that there is some pixelization of the Fourier plane), so the weights are linearly interpolated in ℓ.

We do not have immediately on hand the tools to first combine unpurified maps, and then purify [1] the final ILC map. However we are able to purify individual frequency maps, and do this before implementing the ILC.

The ILC literature derives everything in units of C_ℓ, however the BICEP/Keck pipeline contains many necessary components in units of D_ℓ. Weights calculated in either unit did not produce differing results, so this analysis is done in units of D_ℓ.

3.4 Calculating Weights

This analysis utilizes simulations produced from the BICEP/Keck 2015 (BK15) data release, which consisted of 12 frequency bands: BICEP/Keck frequency maps 95, 150, and 220 GHz, WMAP 23 and 33 GHz, and Planck 30, 44, 70, 100, 143, 217, and 353 GHz. We construct one 12x12 covariance matrix $\mathbf{R}_\ell$ for each ℓ bin from the auto and cross spectra found in the existing pipeline for BK15. We then suppression factor correct the spectra that go into $\mathbf{R}_\ell$, multiplying each element by its corresponding rwf factor ("reciprocal window function", or

---

[1]Purification is the method in which the E to B leakage in the B-mode polarization maps is suppressed.

Figure 3.1: ILC weights calculated for the 17 BICEP/Keck ℓ bins of a single realization for the BK15 (left) and WMAP (right) frequencies. The solid points are the science bins focused upon in this analysis, however weights are calculated in all bins. Note weights in the first bin are zeroed out due to the fact that our scan strategy and TOD filtering remove signal for $\ell < 20$. Weights at ℓ values higher than the 17th bin are also zeroed out for interpolation reasons, as there are no pseudo-C_ℓ calculated in this region.

the inverse of the suppression factor).

$$
\begin{bmatrix}
D_{\ell,(95\times95GHz)} \cdot rwf_{(95\times95GHz)} & \cdots & D_{\ell,(95\times353GHz)} \cdot rwf_{(95\times353GHz)} \\
\vdots & \ddots & \vdots \\
D_{\ell,(95\times353GHz)} \cdot rwf_{(95\times353GHz)} & \cdots & D_{\ell,(353\times353GHz)} \cdot rwf_{(353\times353GHz)}
\end{bmatrix}
\tag{3.14}
$$

Next we calculate one weight per frequency for each ℓ bin using equation (3.11). Our goal is to eventually use this process on real data, and as such for this analysis all weights are calculated using simulations composed of CMB signal, noise, and foregrounds. Weights are calculated per simulation so that each simulation has its own set of weights. Figures 3.1 and 3.2 show weights for individual frequencies in one realization, where the linearly interpolated weights are grouped according to their experiments: the BICEP/Keck frequencies, the WMAP frequencies, and the Planck LFI and HFI frequencies described in section 2.8.
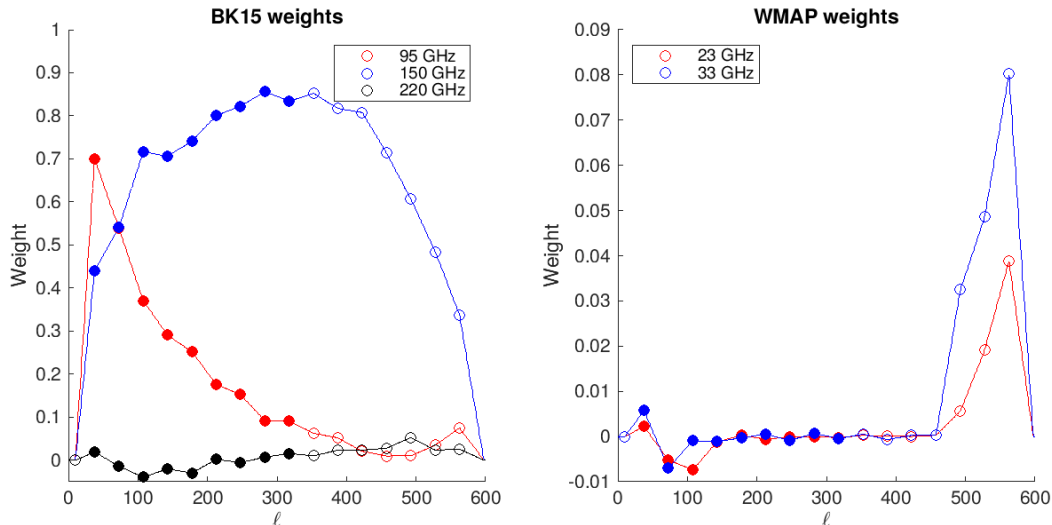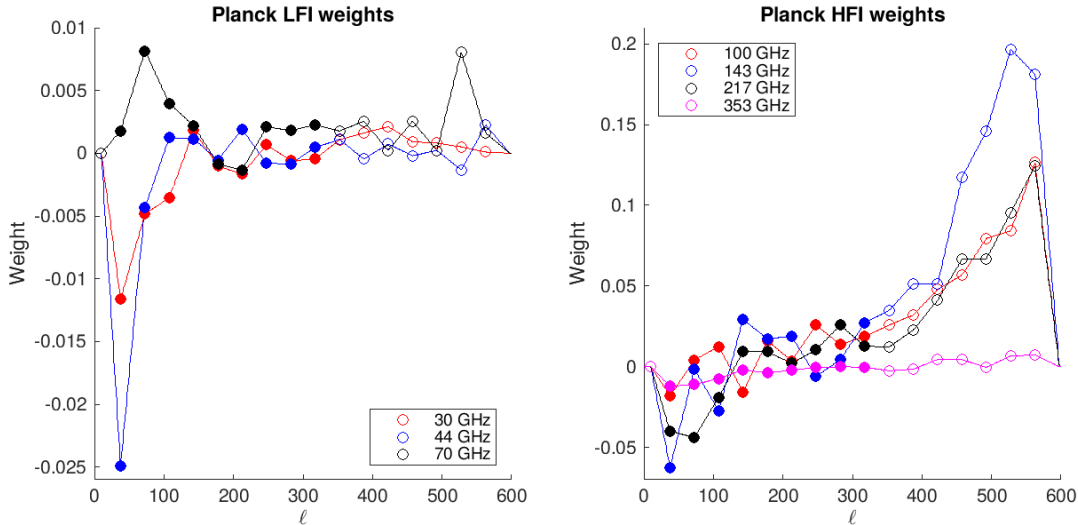
Figure 3.2: ILC weights calculated for the 17 BICEP/Keck ℓ bins of a single realization for the Planck LFI (left) and HFI (right) frequencies. The solid points are the science bins focused upon in this analysis, however weights are calculated in all bins. Note weights in the first bin are zeroed out due to the fact that our scan strategy and TOD filtering remove signal for $\ell < 20$. Weights at ℓ values higher than the 17th bin are also zeroed out for interpolation reasons, as there are no pseudo-C_ℓ calculated in this region.

One can note that some frequencies have very little influence, while others are highly weighted in the ILC map. One effect on the values of the weights for the 95 GHz and 150 GHz BICEP/Keck frequencies is the beam filtering. The 95 GHz instrument beam filtering for BK15 was 43 arcmin FWHM [2] and for 150 GHz 30 arcmin FWHM, meaning that the 95 GHz map has heavier beam filtering than the 150 GHz. We expect then that the 95 GHz map is less sensitive than the 150 GHz map to higher multipoles (which correspond to smaller angular scale), and we see this in the behavior of weights at higher ℓ in the left panel of figure 3.1, where the 95 GHz weights go to zero and the 150 GHz weights remain high. The high frequency weights tend to be negative as they are subtracting foreground signal, and 150 GHz has the highest weights because it is less noisy and drives the signal of the ILC.

As mentioned previously, in practice the weights are linearly interpolated in two dimen-

---

[2]Full width at half maximum (FWHM) is a parameter commonly used to characterize the width of a beam.

sional ℓ space because the maps that they are applied to are continuous in ℓ. In the first and last ℓ bins the weights are forced to zero before interpolation. The filtering process removes signal at $\ell < 20$, and for interpolation reasons the seventeenth bin has zeroed weights because there is no pseudo-C_ℓ calculated for this region of ℓ. The interpolation scheme assumes that the binned weight calculations are done for the center of the bins, for example in the second bin (but first science bin) $ell = 37.5$. This interpolation is done over a matrix containing all of the coordinates (ℓ_x, ℓ_y) of the maps in Fourier space, resulting in a matrix of interpolated weights IW for each frequency.

Observations from the 95 and 150 GHz frequencies in the BK15 weights panel of figure 3.1 contain the lowest noise, hence their high weighting. The weights in the WMAP weights panel on the right are at the low end of the observed frequency spectrum. We see the weights spike around $\ell \approx 500$ for WMAP 33 GHz, and while plotted is a single realization, this trend occurs for all realizations. This may be due to the way in which BICEP/Keck reobserves the WMAP and Planck maps, however because this analysis focuses on the power at much lower values of ℓ not affected by this weighting, this behavior is not explored.

The Planck weights in figure 3.2 include frequencies at the very edges of the observed frequency spectrum where occurs the peak brightness for synchrotron radiation and dust. Because these weights are calculated in CMB units, they tell us how much each map contributes CMB signal to the ILC map rather than how much dust or synchrotron is contributed. Although the weights are quite small compared to the BK15 weights they nonetheless carry an important contribution to the cancellation of foregrounds in the ILC map.

3.4.1   Alternate Interpolation Methods

We explored two other interpolation methods besides a linear interpolation: spline and step. The spline interpolation is a piece-wise cubic polynomial while the step interpolation assumes a step function in ℓ, where all weights for each ℓ bin are the same value. Figures 3.3 and 3.4
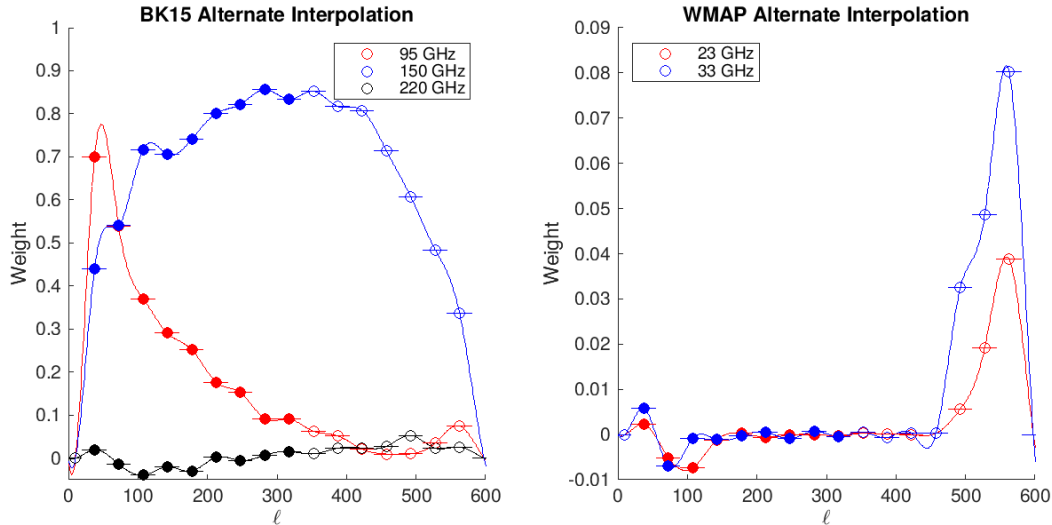
Figure 3.3: ILC weights calculated for the 17 BICEP/Keck ℓ bins of the same realization plotted in figure 3.1. To the left are plotted the weights for the BK15 frequencies, and on the right are the weights for WMAP frequencies. The solid points are the science bins focused upon in this analysis. Note weights in the first bin are zeroed out due to the fact that our scan strategy and TOD filtering remove signal for $\ell < 20$. Weights at ℓ values higher than the 17th bin are also zeroed out for interpolation reasons, as there are no pseudo-C_ℓ calculated in this region. The smooth line denotes a spline interpolation method in the weights, while the horizontal lines show the step interpolation.

show the same weights as the circle points from figures 3.1 and 3.2 calculated for each ℓ bin, this time with continuous lines denoting the spline interpolation and horizontal denoting step interpolation.

Upon examining the output ILC power spectra calculated with weights interpolated with spline or step methods versus a linear method, we found that while the mean of sims (described in section 3.6) was not greatly affected by the method of interpolation, the error on the mean was slightly higher for spline and step interpolation as compared with the linear interpolation. Therefore we chose to continue the analysis using only a linear interpolation on the weights.
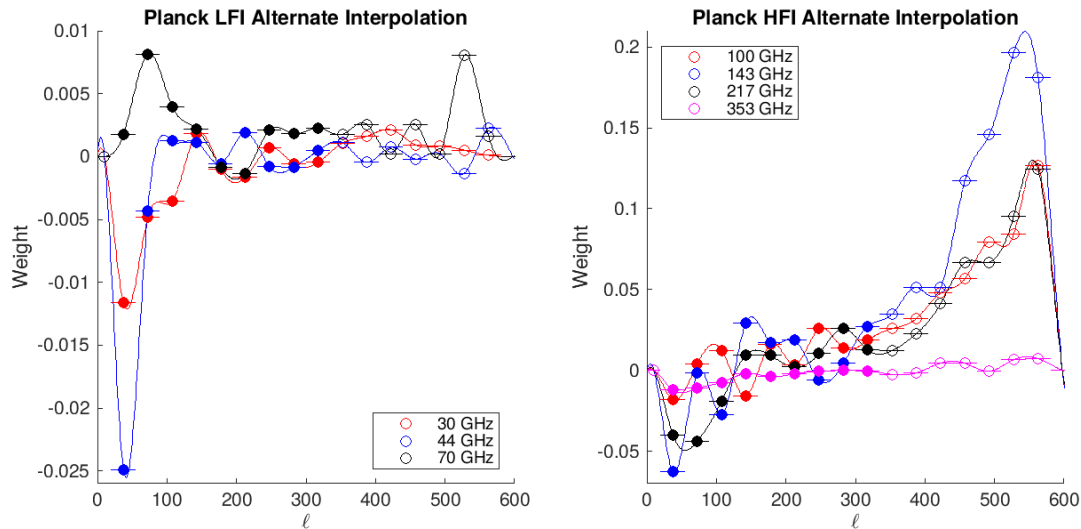
Figure 3.4: ILC weights calculated for the 17 BICEP/Keck ℓ bins of the same realization plotted in figure 3.2. To the left are plotted the weights for the Planck LFI frequencies, and on the right are the weights for Planck HFI frequencies. The solid points are the science bins focused upon in this analysis. Note weights in the first bin are zeroed out due to the fact that our scan strategy and TOD filtering remove signal for $\ell < 20$. Weights at ℓ values higher than the 17th bin are also zeroed out for interpolation reasons, as there are no pseudo-C_ℓ calculated in this region. The smooth line denotes a spline interpolation method in the weights, while the horizontal lines show the step interpolation.

3.5   ILC Map

Our Fourier ILC map is calculated using a modification of equation (3.6):

$$m_{ILC} = \sum_i IW_i \cdot m_i \tag{3.15}$$

where m_i is the purified B-mode Fourier space map, IW_i is the interpolated weight matrix, and i denotes the frequency. While this analysis focuses on B-modes, this process can easily be modified for both E-mode and temperature maps.

Figure 3.5 shows maps in order from top to bottom on the left-hand side (with RA and Dec plotted on x and y axes) the 150 GHz B-mode map from one of the realizations in the BK15 dataset, the ILC B-mode map, and the difference between the two maps. Because the 150 GHz map is by far the highest weighted map in the ILC process, it is most interesting for us to compare directly to the ILC map. Comparing first the top two maps on the left, we see that while large scale structure is preserved in the ILC map, the ILC seems to remove a significant amount of the noisy small structure. The difference map shown in the bottom left panel shows us the amount of noise that is removed by the ILC when compared to just the 150 GHz map. On the right-hand side we have the CMB B-mode map, and below it the 150 GHz dust map. We can see some common structure between the ILC map and the CMB map, and more obviously between the dust map and the difference between the 150 GHz and ILC maps. These CMB and dust maps are useful for comparison because as described in section 3.2, the ILC map is the estimated CMB signal. We expect the ILC map to be more noisy than the CMB because of the leftover noise and foreground residuals. The difference between the 150 GHz and ILC maps should contain the 150 GHz dust signal, which we can easily see is the case when comparing to the dust map on the right.
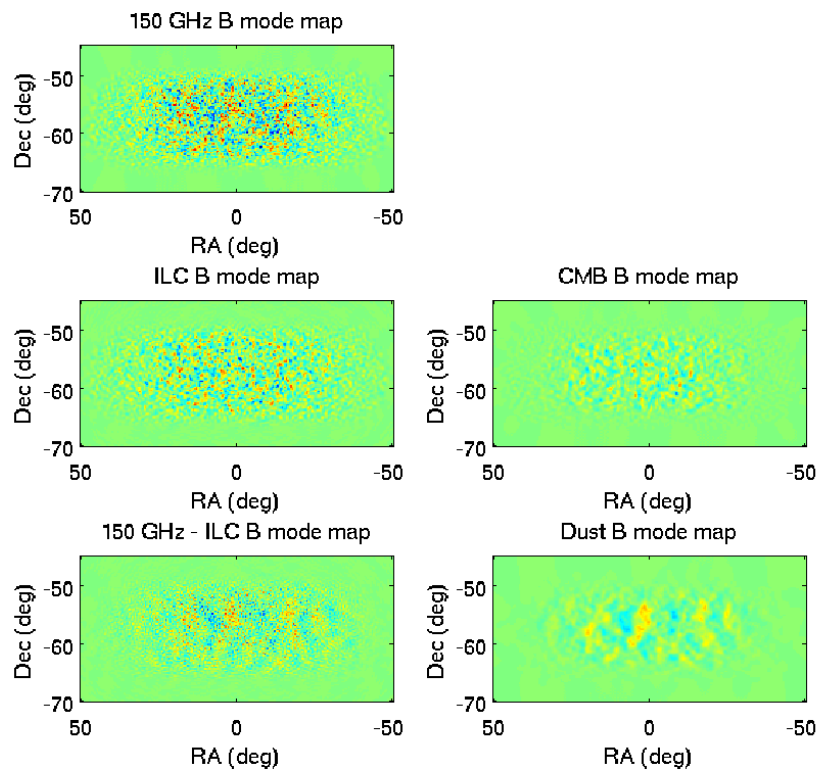
Figure 3.5: Left top: BK15 150 GHz B-mode map. Left center: ILC B-mode map. Left bottom: Difference between ILC and BK15 150 GHz B-mode maps. Right top: CMB B-mode map. Right bottom: BK15 150 GHz dust map.

3.6   ILC Power Spectrum

Now that we have an ILC map, we can calculate the ILC auto power spectrum which also needs to be suppression factor corrected. The rwf factors for the ILC are calculated from the existing frequency rwf factors described in section 3.4. While the calculations for this analysis use 12 maps, a description of the ILC rwf calculation can be easily written out with a two map system:

$$m_{ILC} = W_1 m_1 + W_2 m_2 \qquad (3.16)$$

where W_1 is the weight matrix for map m_1 and W_2 is the weight matrix for map m_2. We can then write the binned uncorrected ILC power spectrum as $m_{ILC} m_{ILC}^*$, where m_{ILC}^* is the conjugate ILC map, and sum this value over the annuli in Fourier space, similar to figure 2.7 in chapter 2.

$$D_{uncorrected,\ell} = \sum (W_1^2 m_1 m_1^* + W_2^2 m_2 m_2^* + W_1 W_2 m_1 m_2 + W_1 W_2 m_1^* m_2^*) \qquad (3.17)$$

We want to correct this power spectrum with the ILC rwf calculated in equation (3.18), where each input frequency auto and cross spectrum has its own corresponding rwf factor.

$$rwf_{ILC,\ell} = W_1^2 rwf_1 + W_2^2 rwf_2 + 2 W_1 W_2 rwf_{1\times 2} \qquad (3.18)$$

rwf_1 and rwf_2 correspond to m_1 and m_2 auto spectra, and $rwf_{1\times 2}$ corresponds to the cross spectra between m_1 and m_2. This equation can be used for any arbitrary number of maps, and is modified to account for the 12 maps that we use. The ILC power spectrum is corrected for each bin with equation (3.19).
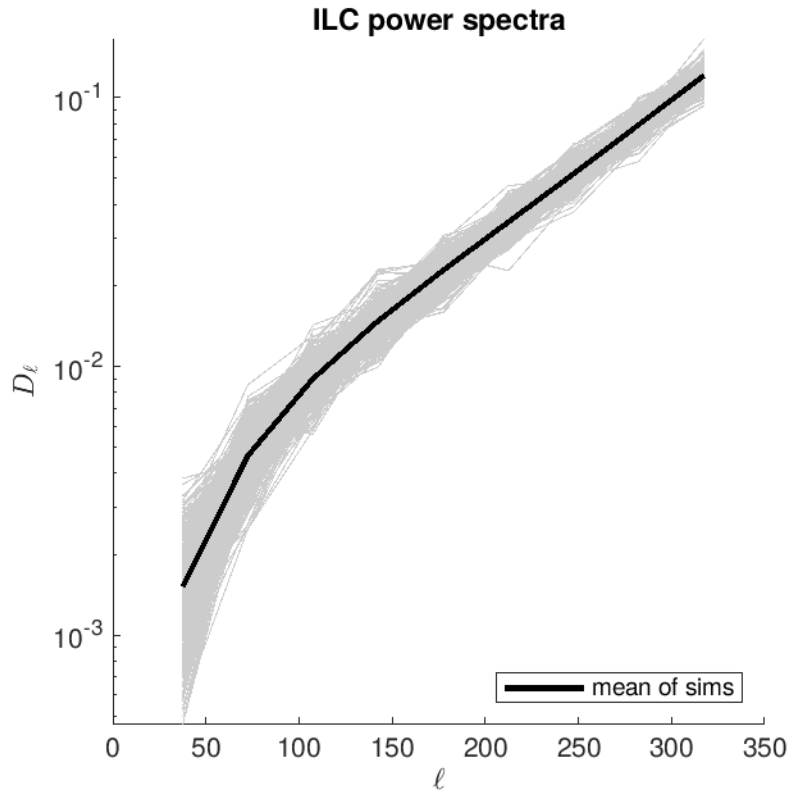
Figure 3.6: Power spectra for the ILC simulations, where grey lines are the individual simulations and the black line is the average of the grey. The power spectra are in units of D_ℓ and calculated for the 9 science ℓ bins.

$$D_{corrected,\ell} = D_{uncorrected,\ell} \cdot rwf_{ILC,\ell} \tag{3.19}$$

The grey lines in figure 3.6 show the results of equation (3.19) for each of the 499 CMB+dust+noise BK15 simulations.

We will describe in the next chapter a derivation of an ILC model which requires ILC simulations to be calculated for the different signal components that go into the sky simulations used to obtain the ILC power spectra in figure 3.6. The sky map simulations are for 499 realizations of the real sky, and each simulation has a corresponding noise only, CMB only and dust only map. We can use these to create six different types of spectra: noise, dust, and CMB only auto-spectra, and the three cross spectra between each of these components
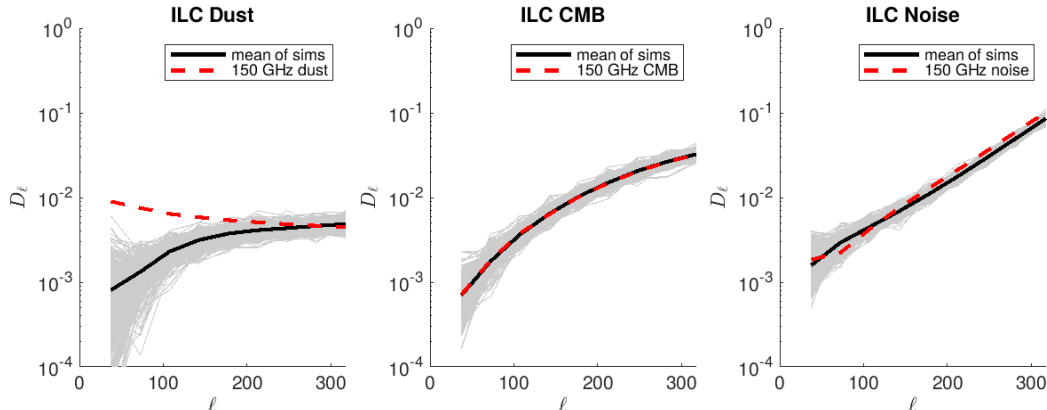
Figure 3.7: Power spectra for the ILC noise, dust, and CMB only simulations, where grey lines are the individual simulations and the black line is the average of the grey. On the left is the dust auto power spectra, center is the CMB power spectra, and on the right is the noise power spectra. The power spectra are in units of D_ℓ and calculated for the 9 science ℓ bins. The means of the corresponding BK15 150 GHz simulations are plotted in dashed red.

(dust $\times$ CMB, dust $\times$ noise, and noise $\times$ CMB).

We calculate the auto spectra using weights obtained from the CMB+dust+noise maps and apply to the map of interest to create a dust/CMB/noise only map, and then calculate the auto spectrum of this ILC map in the same manner as was done for the CMB+dust+noise ILC map. We can directly compare these APS to those of the BK15 150 GHz dust/CMB/noise only simulations as an additional check to the map to map comparison shown in figure 3.5. Figure 3.7 shows the auto power spectra for each of these simulation types, with dust APS on the left, CMB APS in the center, and noise on the right. The grey lines are the individual simulations, the black lines are the means of the grey, and the dashed red are the means of the corresponding BK15 150 GHz simulations.

All panels are plotted on the same logarithmic y-axis, and we can see immediately when comparing the ILC dust APS to the average of 150 GHz dust simulations that we have removed dust on average by an order of magnitude at the lowest science bin. We can see in the center panel that we preserve the CMB signals through our process, and in the right panel that for $\ell > 100$ we have a consistent level of noise.

Figure 3.8 shows the cross terms between the dust, noise, and CMB only maps. In the
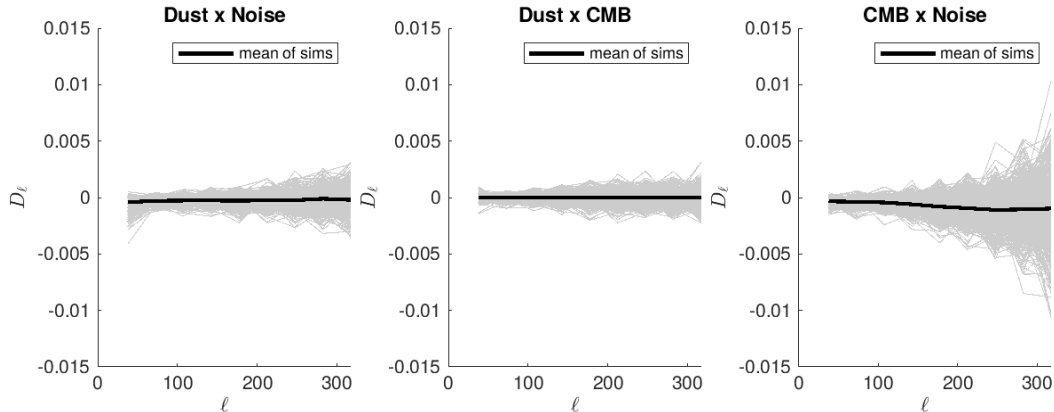
Figure 3.8: Cross terms between the ILC CMB, noise, and dust simulations. The limits for the y-axis are the same for all three panels. On the left are the dust×noise cross terms, the center panel shows the dust×CMB cross terms, and on the right are the CMB×noise cross terms. The grey lines denote the individual realizations, and the black lines are the average of the grey.

right panel are the dust×noise cross terms, in the center panel are the dust×CMB cross terms, and on the right are the CMB×noise cross terms. The cross terms for individual realizations are plotted in grey, and the average of the grey realizations is plotted in black.

These cross terms are calculated by applying the same weights that have been used thus far to the dust/CMB/noise only map simulations, and then the cross spectra between the different types of maps can be calculated in the same manner as an auto cross spectrum being the cross of a map by itself. On average, because CMB, dust and noise are three uncorrelated sources, we would expect their cross terms to be zero. We can see most obviously from the cross terms plotted in the right panel that this is not the case, the reasons for which are described further in section 4.1.1. We can see more clearly the average cross term divergence from zero in figure 4.3.

3.7 CMB Component

The CMB component that we wish to compare to the ILC power spectrum is calculated by applying an ILC weighted BPWF (described in section 2.7) to the lensing CMB theory

spectrum (which includes the lensing B-modes), as detailed below.

The ILC weights are applied to the BPWF of each frequency in a similar manner as they were applied to maps in equation (3.15). However, since the BPWF are one dimensional in ℓ, they are multiplied by a vector of weights interpolated along ℓ rather than a matrix interpolated in the 2D Fourier space of (ℓ_x, ℓ_y). The equation for the ILC BPWf is

$$bpwf_{ILC}(\ell) = \sum_{i,j} iw_i \cdot iw_j \cdot bpwf_{i,j} \qquad (3.20)$$

where $bpwf_{i,j}$ is the band power window function for frequencies i, j and iw is the interpolated weight vector for frequency i, j. As mentioned previously, historically the BICEP/Keck collaboration bins the power spectra into 17 bins of width 35 multipoles. These estimated bandpowers, however, are not simply the mean over the bin width, rather are sensitive to adjacent bins. The BPWF provide the contribution of each multipole to the bandpower of interest. Figure 3.9 shows the ILC BPWF compared to the 150 GHz BPWF for all bandpowers. We have seen that the 150 GHz map is highly weighted in this ILC, which is why the ILC BPWF follows the 150 GHz BPWF very closely for most of the bandpowers [3]. We see through the local maxima in this plot the contribution of other multipoles to each bandpower.

Once we have our ILC BPWF, we can then calculate the CMB component for each ℓ bin:

$$\langle \hat{D}_\ell^c \rangle = \sum_\ell bpwf_{\ell,ILC} \cdot D_{\ell,LCDM} \qquad (3.21)$$

where $D_{\ell,LCDM}$ is the scaled ΛCDM theory spectrum for the ILC where r can be tuned to the desired value. $bpwf_{\ell,ILC}$ are the ILC BPWF from equation (3.20).

Figure 3.10 compares the CMB component (magenta) for $r = 0$ to the CMB only ILC realizations (grey) with their mean plotted as a solid black line. These are the same CMB

---

[3]The BPWF for all other frequencies are very similar to the 150 GHz BPWF.
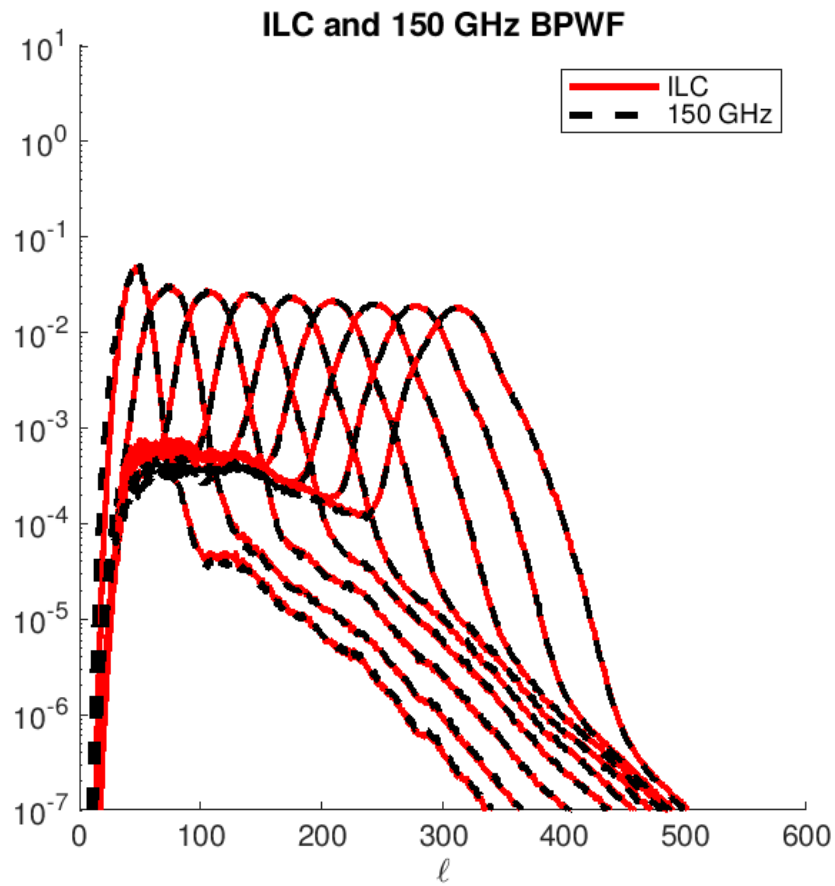
Figure 3.9: Bandpower window functions for BK15 150 GHz (dashed black) and ILC (red) power spectra. We see that the ILC BWPF follow very closely the 150 GHz BPWF, as this frequency has the highest weighting.
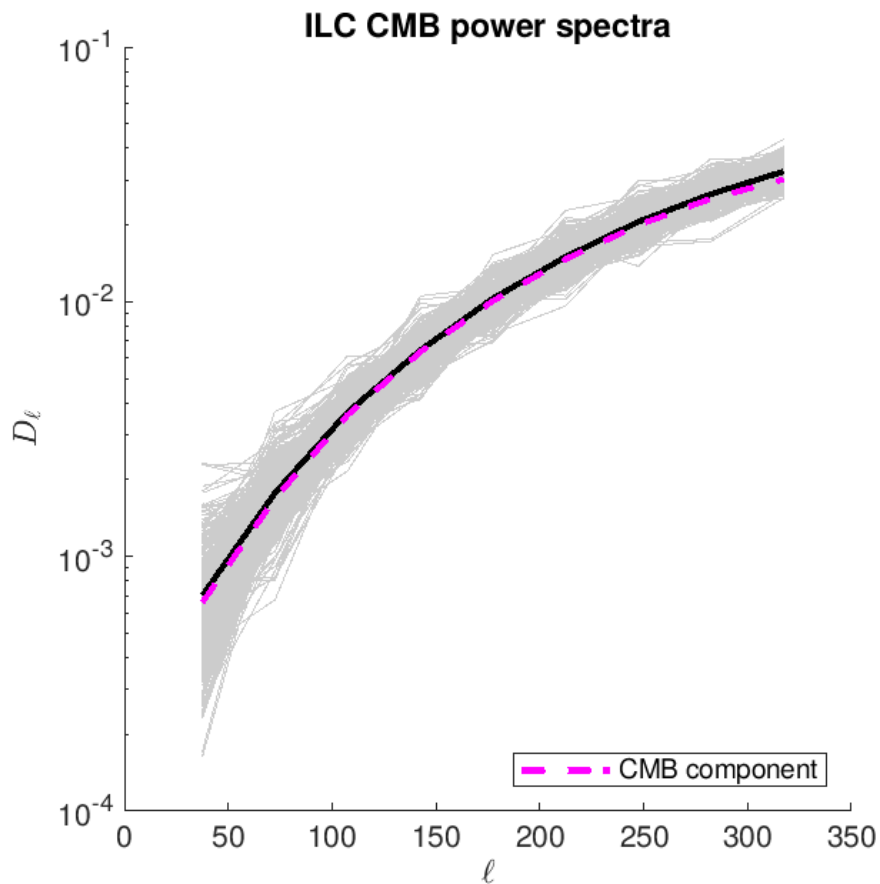
Figure 3.10: The grey lines are the ILC CMB simulations, which have been calculated by applying weights obtained from full sky simulations and to CMB only simulations. The black line is the average of the grey, and the CMB component is plotted in the dashed magenta line.

ILC realizations plotted in the center panel of figure 3.7. There is clearly a strong agreement between the black and dashed magenta lines, which confirms that we are calculating our ILC rwf values correctly. The normalization factor of the BPWF is the suppression factor, for which the rwf is the inverse. Recall from section 3.6 the ILC power spectra is corrected with the ILC rwf factors, and our CMB component is calculated with ILC BPWF calculated in a similar way to the rwf, the agreement between which confirms the calculations are done correctly.

Chapter 4

Likelihood Analysis

Here I present the likelihood analyses performed with two different ILC models. The first is a simulation based model, for which we conduct a multi-parameter maximum likelihood search (ML search) over CMB and dust parameters. The second model we explore is the analytic Saha ILC model for which we do a one dimensional ML search in r space, having first marginalized over foreground parameter space. The results of an ML search provide the values of the specified parameters for which the data has the highest probability.

The parameters used in this analysis are as follows, recalling that dust and synchrotron signals follow a power law in both ℓ [1] and frequency:

- A_d: the amplitude of dust power. We specify this amplitude for $\ell = 80$ at the frequency 353 GHz, or it's "pivot" frequency. This is the frequency for which we could change the scale factor β_d without affecting A_d.

- β_d: the spectral index scaling factor describing how dust scales with frequency.

- A_s: the amplitude of synchrotron power. This amplitude is specified at $\ell = 80$ for the

---

[1] We consider only the frequency scaling. The descriptor for how dust(synchrotron) scales with ℓ is $\alpha_d(\alpha_s)$. We choose to define the dust and synchrotron parameters at $\ell = 80$ because we expect the CMB signal to be strongest here.

same reason as A_d, however the pivot frequency for this parameter is 23 GHz.

- β_s: the spectral index scaling factor describing how synchrotron scales with frequency.

- r: the tensor-to-scalar power ratio.

4.1 ILC Models

We run our likelihood analysis in a simplified framework, using fewer free parameters than the standard BICEP/Keck likelihood analysis. This section will detail the two models that were tested. First I will describe a simulations based model using the realizations from BK15, and following that will be an analytic model derived by Rajib Saha in his 2008 paper [36].

4.1.1 Simulation Based ILC Model

A large part of this analysis includes studying the effects of varying the dust model. We create a modified sky simulation by adding simulations with CMB to those with noise and with dust (as opposed to the BICEP/Keck sky simulation, which has all of these signal included from the start). For this model we did not consider synchrotron for reasons discussed further in section 4.1.2. The simulations containing only dust can be scaled with respect to the A_d and β_d parameters with a scale factor calculated relative to the dust in the 150 GHz map with an internal BICEP/Keck pipeline function specifically written for frequency scaling. This function is multiplied by another scale factor dependent on the variable A_d, all of which is done before the maps go through the ILC process. We create ILC maps with these sky simulations as well as for dust, CMB, and noise only simulations which are used for modeling purposes. The calculation of the ILC power spectra is done in the same manner as described in section 3.6.

Due to expensive computation time the simulated ILC power spectra are only created for

discrete values in a limited range of A_d and β_d values where $r = 0$. To conduct a likelihood analysis, we create a model which evaluates the power spectra for arbitrary values of these parameters:

$$ILC_{aps} = \langle D_\ell^c \rangle + b_{noi} + b_{ILC} + b_{fg} \tag{4.1}$$

where $\langle D_\ell^c \rangle$ is the contribution to the ILC power spectrum by the CMB signal from equation (3.21). b_{noi} is the noise contribution to the ILC power spectrum, or the noise bias, calculated from noise only simulations combined together with ILC weights calculated from full sky simulations. b_{ILC} is the ILC bias contribution from dust × CMB and CMB × noise cross terms. Because of the way the ILC is applied to each realization individually there are some chance correlations that appear between these cross terms where we would expect there to be none. b_{fg} is the residual foreground bias, containing contributions from dust power that has not been completely removed in the ILC process.

Noise Bias

We have a good understanding of the noise contribution from instruments and atmosphere and are able to calculate a noise bias by applying our sky ILC weights to noise only simulations, however even in the ideal case of zero foregrounds we still could not completely remove noise and there will always be a non-zero contribution to the ILC power spectrum. While the noise simulations have no dependence on dust parameters A_d and β_d, the overall noise bias will have some small dependence because the ILC weights are calculated with a full sky map which does contain dust. To account for this a noise bias must be calculated for arbitrary values of A_d and β_d for this model.

Our ILC weights are calculated from the sky simulations for discreet values of A_d and β_d spanning $\beta_d = [1.2, 2]$ and $A_d = [0, 10]$, and can obtain an ILC noise map by applying

these weights to noise only simulations, and from that we can calculate an ILC noise power spectrum. This is repeated for all 499 realizations, and the average of these simulations is our noise bias, which is plotted as the green line in figure 4.1 for the nominal values $A_d = 3.75, \beta_d = 1.6$. We note that this curve closely follows the black line, the average of full sky ILC power spectra, and can see that the dominant contribution to the cleaned power spectrum is noise. We perform a two dimensional linear interpolation of the noise bias in the (A_d, β_d) parameter space for each individual ℓ bin, and show the surfaces for each science bin in figure 4.2. We can easily see from the shape of these surfaces the dependence of the noise bias on A_d and β_d.

Foreground Bias

As mentioned previously, our ILC power spectrum still contains a foreground contribution, call the residual foreground bias. We do not have a strong understanding for the behavior of the galactic foregrounds like we do with the noise, so cannot predict how the residual foregrounds will behave once the ILC has been implemented. To get around this, we model the foreground residuals using the foreground power spectra and the cross spectra between ILC noise and foreground maps. As mentioned earlier for the purpose of this section we only consider galactic dust, assuming no synchrotron.

Figure 4.3 shows the average of the cross spectra between noise and dust for all realizations simulations in yellow. As dust and noise are two uncorrelated objects, one would expect that on average their cross spectra would be zero. Instead, we observe a negative trend in their average due to chance correlations between the noise and dust, where there are some simulations in which the dust will randomly cancel with some of the noise in the maps. Ordinarily the ILC would want to remove this dust, but in these cases will leave it alone. Because the ILC weights are calculated per realization, on average this will look like there is some negative correlation between dust and noise, and must be accounted for when modeling
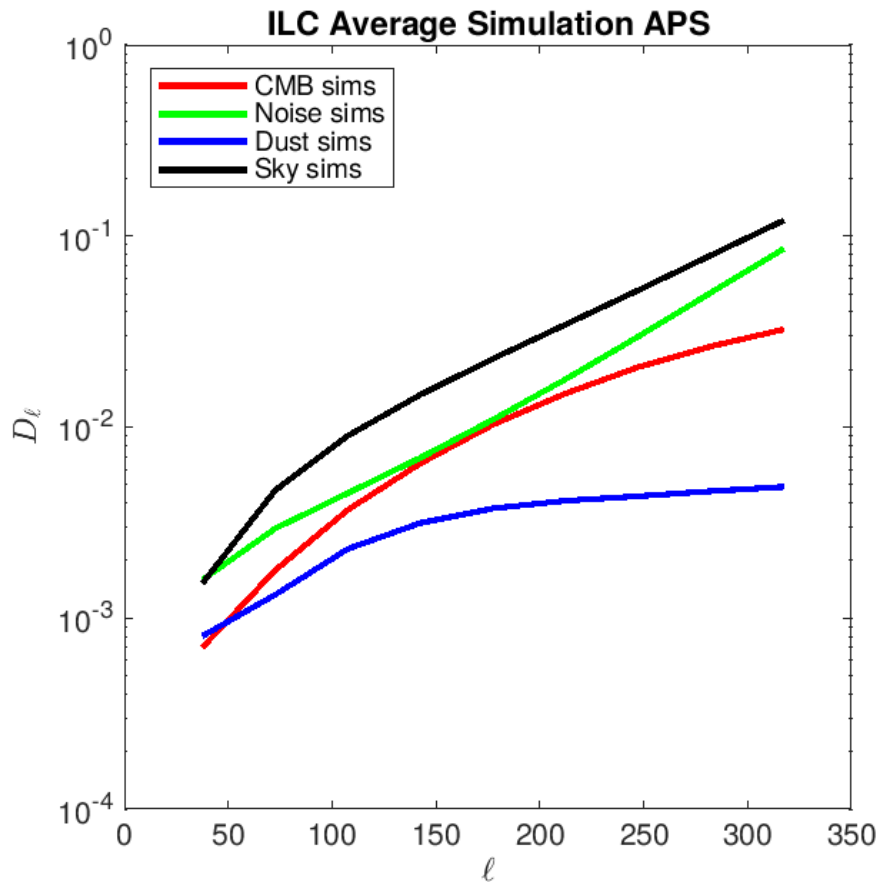
Figure 4.1: Average power spectra of ILC simulations. The black line shows the average of the sky ILC power spectra, green shows the average of noise only ILC power spectra, red shows the average of the CMB ILC power spectra, and the blue shows the average of dust ILC power spectra. Each of the different types of power spectra are calculated with weights obtained from full sky simulations.

Figure 4.2: Noise bias surfaces for each science ℓ bin. The science bins are bins 2-10 of the 17 total ℓ bins described in section 2.7, and are ordered above going from left to right where bin 2 is shown in the top left, to bin 10 on the bottom right. The titles of each surface are the ℓ values at which the bins are centered.

the ILC power spectrum. The blue line in figure 4.1 shows the power spectrum average over all ILC dust simulations with $A_d = 3.75$ and $\beta_d = 1.6$. We can see that the contribution of dust to the full ILC power spectrum is quite small compared to the noise bias, however is still a positive one. The way we calculate our foreground residual bias is

$$b_{fg} = \langle D_{\ell,d} \rangle + 2\langle D_{\ell,d \times n} \rangle \tag{4.2}$$

where $\langle D_{\ell,d} \rangle$ is the average ILC dust only power spectrum over all simulations, and $D_{\ell,d \times n}$ is the average cross power spectrum between ILC dust and noise maps over all simulations.

Again the curves shown in figure 4.3 show results for weights and simulations with nominal values of $A_d = 3.75$ and $\beta_d = 1.6$. We create simulations for our normal range of β_d and A_d, and perform a two dimensional surface fit over dust parameter space for each ℓ bin in the same manner as the noise bias calculation. Figure 4.4 shows the foreground residaul bias surface for each of the science ℓ bins. We find that the behavior of dust as a function of A_d and β_d is different for the first two bins than in the rest, especially in the first bin which is centered at $\ell = 37.5$, where for positive values of A_d this bias actually goes negative. This is because at low ℓ, the negative cross terms between dust and noise overpower the dust only power spectrum, which can be seen visually when comparing the blue line in figure 4.1 to the yellow line in figure 4.3.

ILC Bias

The ILC bias is the contribution of the CMB $\times$ noise and CMB $\times$ dust cross terms to the ILC power spectra. We ignore the dust $\times$ noise term since that is included in the foreground bias calculation above.

For this calculation, we look to the blue and orange lines in figure 4.3. Because the noise level in the maps is much larger than dust and CMB, the CMB $\times$ noise term dominates the
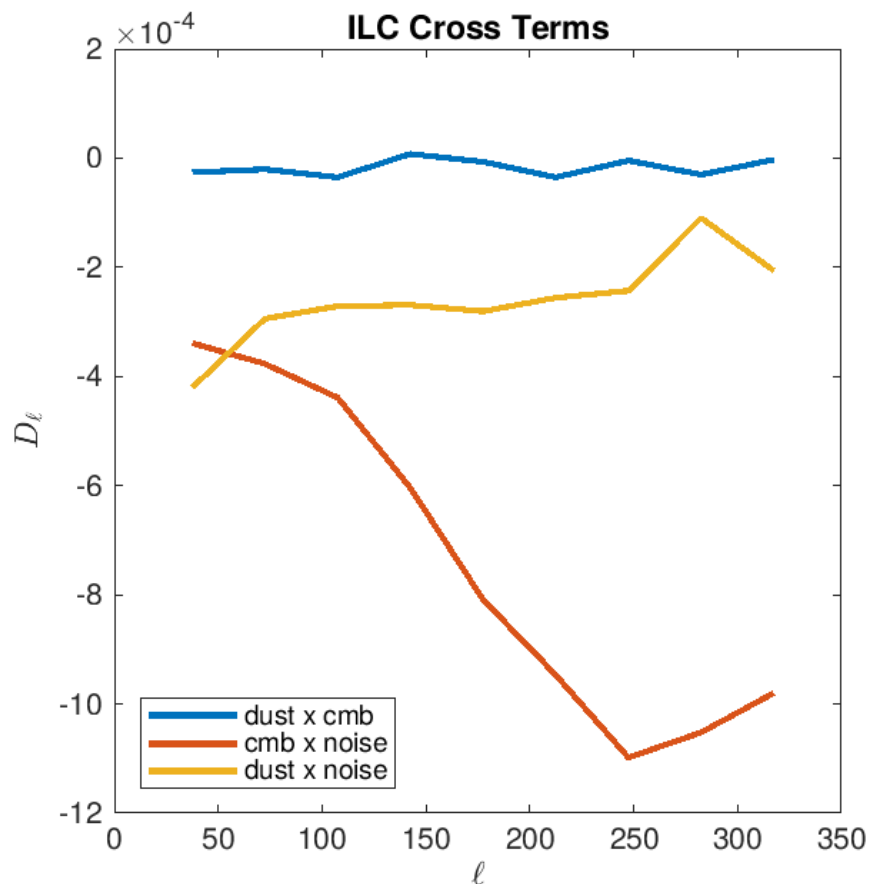
Figure 4.3: Average cross power spectra of ILC simulations. The blue line is the average of the dust $\times$ CMB simulations, yellow is the average of the dust $\times$ noise simulations, and the oragne line is the average of the CMB $\times$ noise simulations. While the fact that the three signals (CMB, noise and dust) are uncorrelated would lead us to believe the average of these cross terms would be zero, we actually find that in fact there is a negative trend.
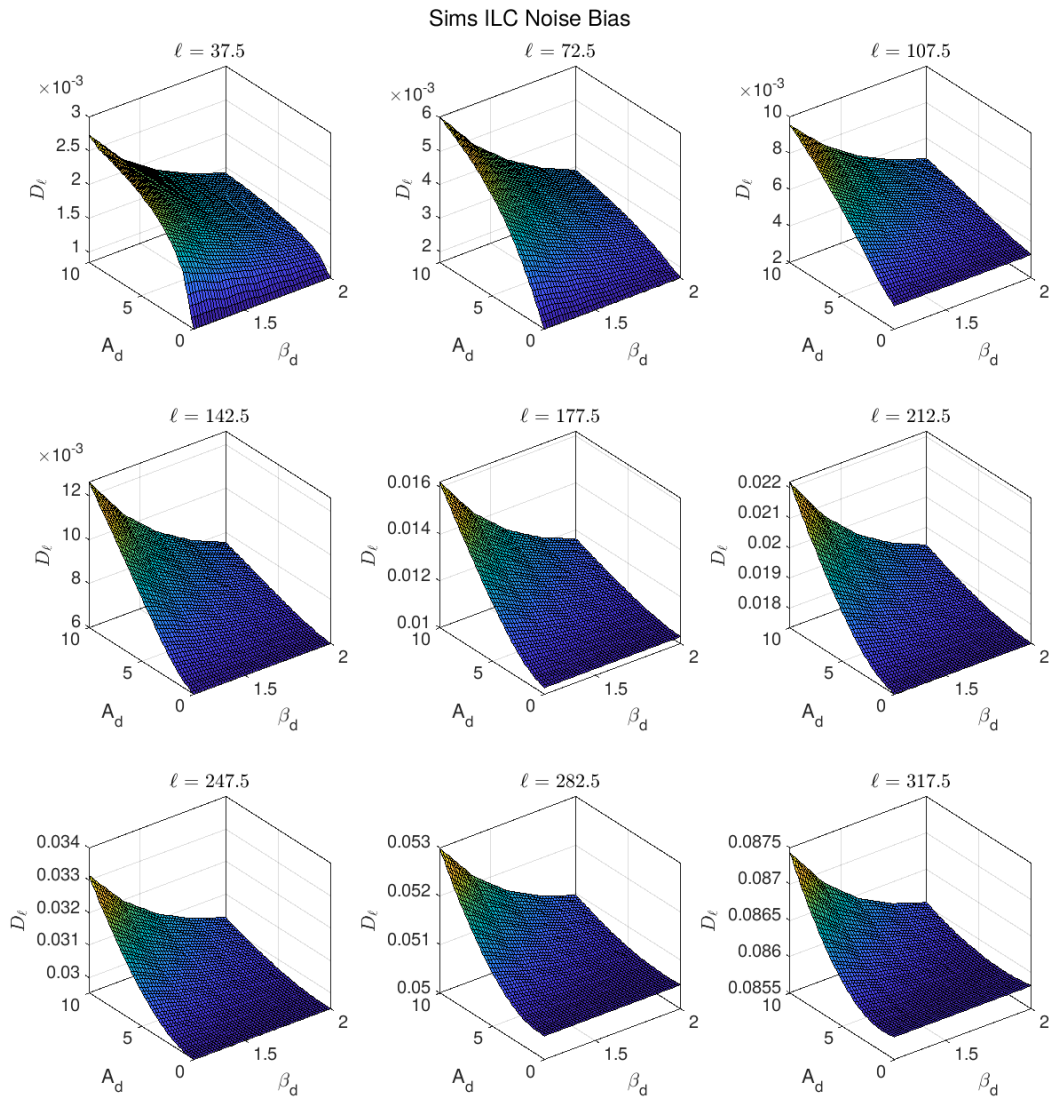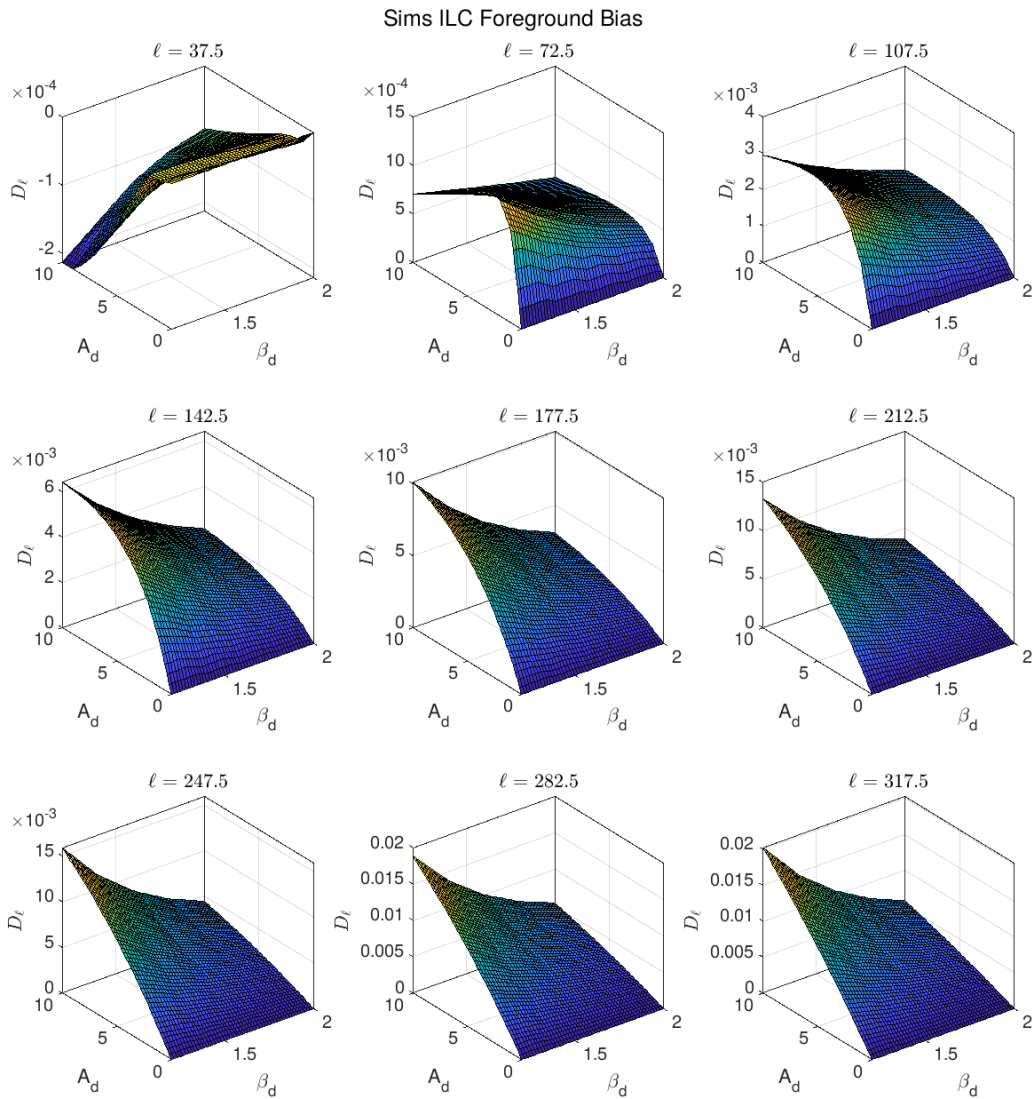
Figure 4.4: Foreground bias surface fits for each science ℓ bin. The science bins are bins 2-10 of the 17 total ℓ bins described in section 2.7, and are ordered above going from left to right where bin 2 is shown in the top left, to bin 10 on the bottom right. The titles of each surface are the ℓ values at which the bins are centered.

dust $\times$ CMB term, and as a result the ILC bias does not change significantly if we vary the values of A_d and β_d, even though both the CMB and noise ILC maps are calculated with full sky weights. The fact that the ILC bias is negative tells us that a small fraction of the true signal is lost when performing the ILC. Figure 4.6 shows in grey the envelope of values the ILC bias takes for different values of A_d and β_d, with the solid black line showing the bias for nominal values of $A_d = 3.75$ and $\beta_d = 1.6$. The ILC bias is calculated in equation (4.3):

$$b_{ILC} = 2(\langle D_{\ell,CMB \times n} \rangle + \langle D_{\ell,d \times CMB} \rangle \tag{4.3}$$

where $\langle D_{\ell,CMB \times n} \rangle$ is the average of the CMB $\times$ noise ILC cross terms over all realizations, and $\langle D_{\ell,d \times CMB} \rangle$ is the average of the dust $\times$ CMB terms. In the same manner as the noise and foreground bias previously described, we calculate these terms for the same range of A_d and β_d values and fit with a two dimensional linear interpolation to obtain values for arbitrary A_d and β_d. Figure 4.5 shows the surface of the ILC bias in each of the science ℓ bins.

4.1.2 Simulation Based ILC Model Result

The results for the simulation based ILC model from equation (4.1) are shown in figure 4.7 as the red dashed line. We can see that the model agrees very well with the average of simulations, and we will see in section 4.4 that this model yields promising results when implemented in a ML search. However, while this model works very well for the dust only case, we do not explore further results for the case assuming synchrotron is present in the maps. This is because this method relies heavily on creating simulations for many values of the dust and synchrotron parameters. We see already that dust has two parameters (A_d and β_d), and adding synchrotron increases our free parameters by another two (A_s and β_s), possibly three if we assume some correlation between dust and synchrotron (ϵ).
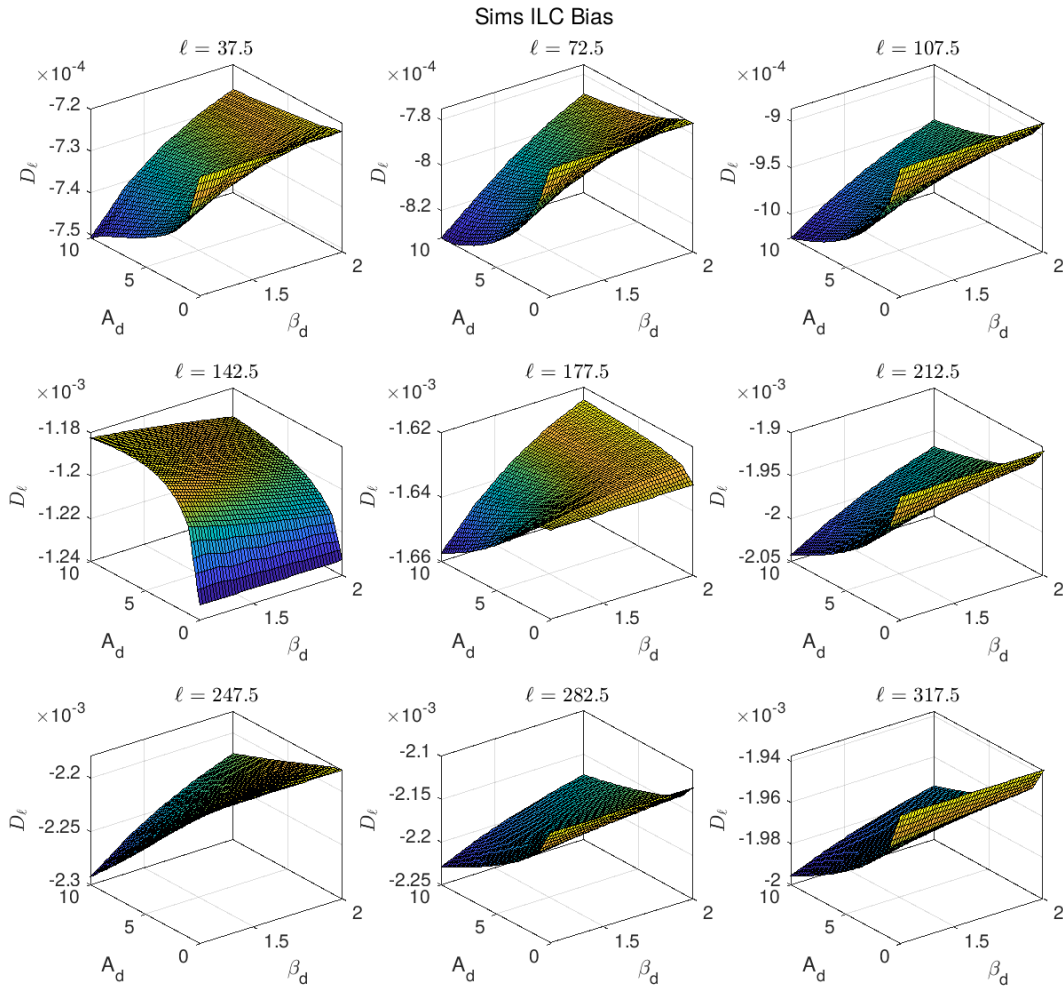
Figure 4.5: ILC bias surface fits for each science ℓ bin. The science bins are bins 2-10 of the 17 total ℓ bins described in section 2.7, and are ordered above going from left to right where bin 2 is shown in the top left, to bin 10 on the bottom right. The titles of each surface are the ℓ values at which the bins are centered.
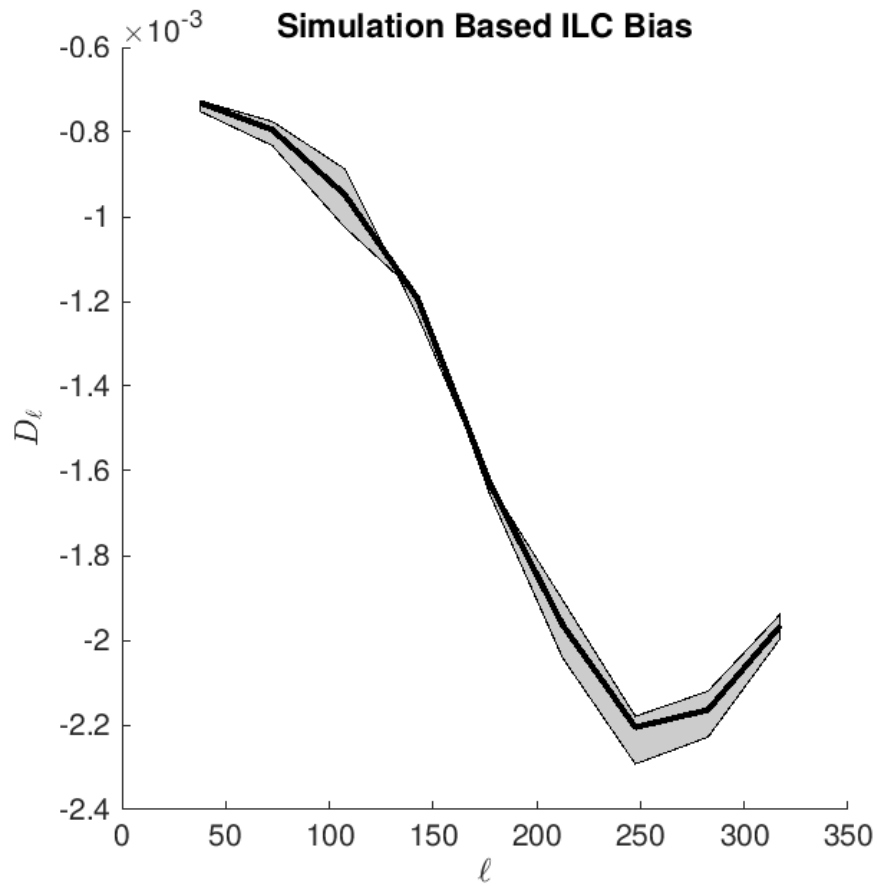
Figure 4.6: ILC bias at 9 science ℓ bins. The black line is calculated for $A_d = 3.75$ and $\beta_d = 1.6$, and the grey area shows the range of values that the bias can take for the range of A_d and β_d that have been calculated.
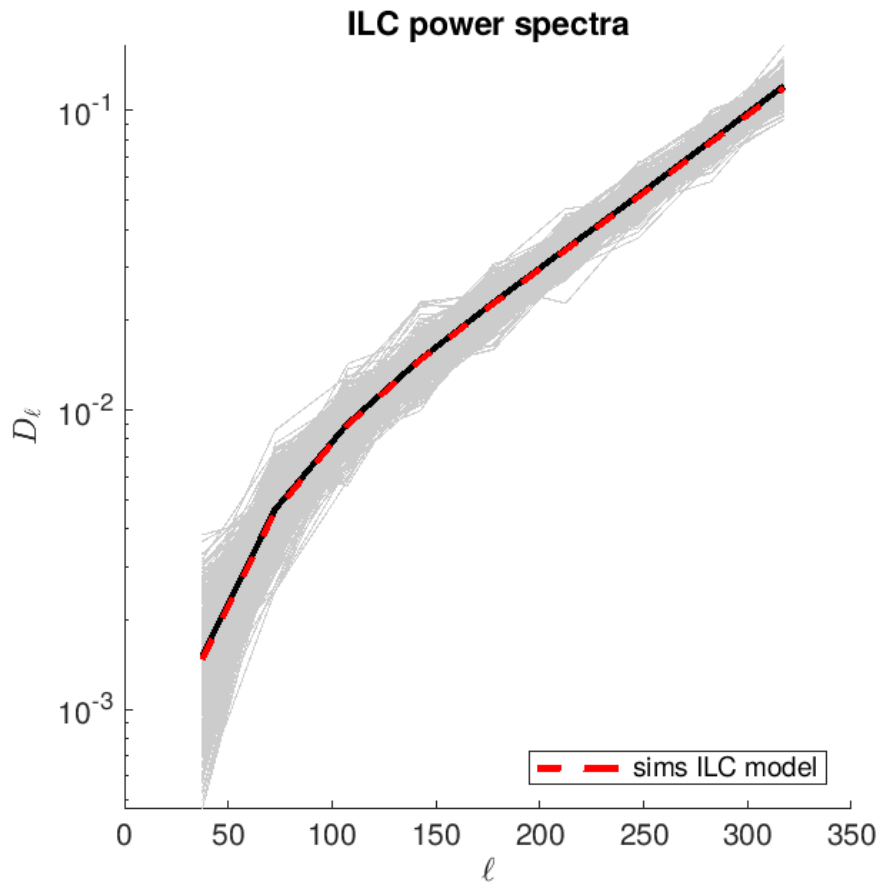
Figure 4.7: ILC power spectra, where grey lines are individual realizations, the black line is the mean of realizations, and the red dashed line is the simulation based ILC model.

Another drawback to this method is because the model is built entirely off of simulations, we need to completely start over should the types of simulations we use change. For example, the simulations shown here are standard simulations where we assume Gaussian dust and no synchrotron. However, there are altdust simulations described in section 4.4 which have different dust and synchrotron models than the standard simulations. It is possible to reproduce this ILC model process with these simulations, however the amount of computation time required to do so complicates things prohibitively.

4.1.3 Saha ILC Model

The Saha ILC model was derived at length by Rajib Saha et al. [36]. This model contains components reflecting the theory lensing CMB signal, noise and foreground residual contributions, and a bias implicit to the ILC method. Rather than working from the map level, this equation forms the power spectrum resulting from the ILC method, shown in equation (4.4). The work done by Saha is in units of C_ℓ, however because my work is done in units of D_ℓ [2] I have slightly modified his equations to reflect this.

$$\langle \hat{D}_\ell^{Clean} \rangle = \langle \hat{D}_\ell^c \rangle + \left\langle \frac{1}{\mathbf{e_0^T}(\hat{\mathbf{D}}_\ell^{\mathbf{f+N}})^{-1}\mathbf{e_0}} \right\rangle + (1 - n_c)\frac{\langle \hat{D}_\ell^c \rangle}{2\ell + 1} \tag{4.4}$$

Here $\langle \hat{D}_\ell^{Clean} \rangle$ is the ILC power spectrum, aiming to reflect the mean ILC simulation based power spectrum (the black line in figure 4.1). $\langle \hat{D}_\ell^c \rangle$ is the same CMB component from the simulation based ILC model in equation (4.1).

The second term in equation (4.4) contains the noise and foreground contributions to the ILC power spectrum, and will for simplicity will be referred to as $S2$. Here $\mathbf{e_0}$ is a column vector containing unit elements, and $\hat{\mathbf{D}}_\ell^{\mathbf{f+N}}$ is a covariance matrix of the noise and foreground frequency maps. The diagonal contains auto power spectra from maps containing only noise and foregrounds, and the off-diagonals are the cross power spectra for these maps.

The last term is the ILC bias, which comes from the chance correlations in the cross spectra described in section 4.1.1. n_c is the number of frequency maps going in to the ILC, which for this analysis is 12. By the very nature of the ILC being a combination of multiple maps, this bias is negative.

---

[2] $D_\ell = \frac{\ell(\ell+1)C_\ell}{2\pi}$

Adjusted Saha Model

The Saha model as written in equation (4.4) does not quite agree with the average ILC simulation power spectra due to the noise and foreground component, as can be seen in the middle panel of figure 4.9. We need to make some changes to the Saha model before proceeding with the likelihood analysis. This adjusted model is shown in equation (4.5).

$$\langle \hat{D}_\ell^{Clean} \rangle = \langle \hat{D}_\ell^c \rangle + \left(\frac{N_{mode} - n_{eff} + 1}{N_{mode}} \right) \frac{1}{\mathbf{e_0^T} \langle \hat{\mathbf{D}}_\ell^{\mathbf{f+N}} \rangle^{-1} \mathbf{e_0}} + (1 - n_c) \frac{\langle \hat{D}_\ell^c \rangle}{N_{mode}} \qquad (4.5)$$

First we note the change in the ILC bias term where the $2\ell+1$ denominator is now N_{mode}, the number of χ^2 degrees of freedom per ℓ bin. This change is due to the small size of the BICEP/Keck observation patch, and because the effective number of pixels in our maps is small at low ℓ due to correlations between the map pixels.

The more subtle change to the $S2$ term is the placement of the brackets around the foreground and noise covariance matrix, as opposed to the whole fraction. We know how to calculate the expectation value of this covariance matrix, where the auto and cross power spectra are obtained from the bandpower expectation values from the BICEP/Keck multi-component model, and can be calculated for arbitrary values of the dust and synchrotron parameters. The expectation value of the entire fraction is another matter that we do not currently have the tools to calculate. Our way of working around this is through the multiplicative factor dependent on N_{mode} and n_{eff}.

We obtained the multiplicative factor $\left(\frac{N_{mode} - n_{eff} + 1}{N_{mode}} \right)$ through the use of a toy ILC model where we assume a case of 4 input maps, all of which only contain independent noise. We created 500 ILC power spectrum realizations and compared the average of the realizations to both $\frac{1}{\mathbf{e_0^T} \langle \hat{\mathbf{D}}_\ell^{\mathbf{f+N}} \rangle^{-1} \mathbf{e_0}}$, and through trial and error found the best results for the scale factor for the modified term $\left(\frac{N_{mode} - n_{eff} + 1}{N_{mode}} \right) \frac{1}{\mathbf{e_0^T} \langle \hat{\mathbf{D}}_\ell^{\mathbf{f+N}} \rangle^{-1} \mathbf{e_0}}$, where for the BK15 dataset we found a best fit for $n_{eff} = 8$. We repeated this process varying the number of modes in the toy maps, and
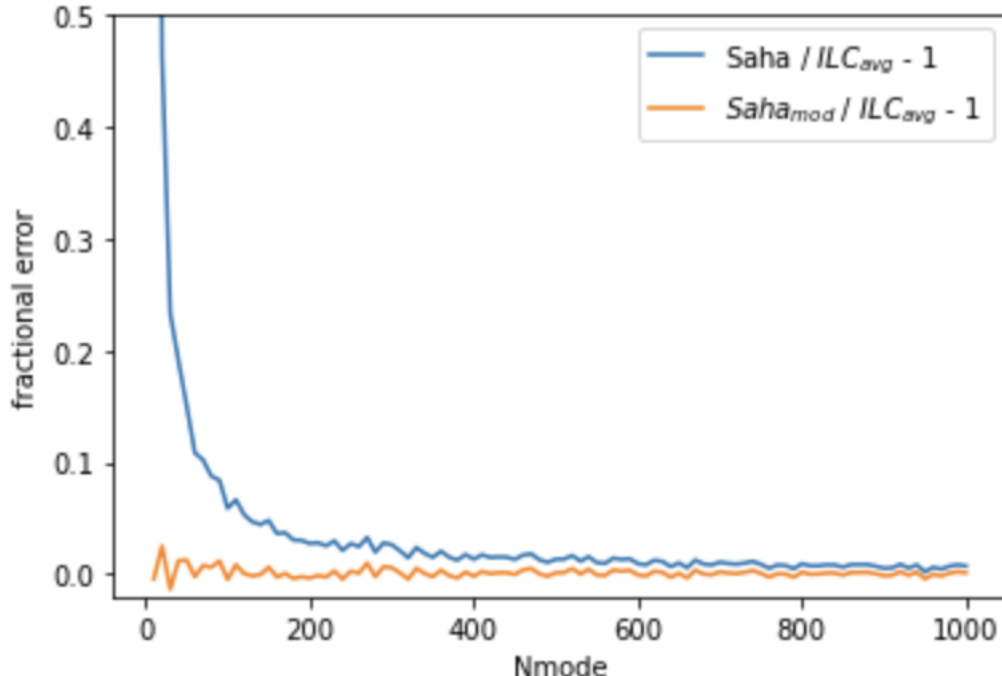
Figure 4.8: Fractional error between ILC toy model and variations on Saha noise term (all maps only contain noise in this toy model). The blue line shows the ILC average compared to the Saha noise term without the multiplicative scale factor, and orange shows with the comparison with the scale factor. The x-axis is number of modes in the maps, and we see that for low number of modes the term without the scale factor becomes quite large in comparison to the ILC average.

plotted the fractional errors in figure 4.8. We can clearly see that without the scale factor, for low number of modes the Saha $S2$ term becomes quite large in comparison to the ILC toy model average, and once we include this scale factor the agreement between the $S2$ term and ILC average greatly improves. This is most obvious at low number modes, however we can see even at higher numbers of modes the orange line is closer to zero than the blue line. Once we obtained this scale factor, it was found empirically that the $S2$ term agrees best with our ILC average using BK15 realizations when the effective number of maps $n_{eff} = 8$.

Figure 4.9 shows the individual components of equation (4.5) compared with the corresponding ILC simulations. Going from left to right, the first panel shows the CMB component, or the first term in equation (4.5), compared with the ILC CMB only simulations where the black dotted line is the mean of grey simulations. The center panel shows the Saha noise
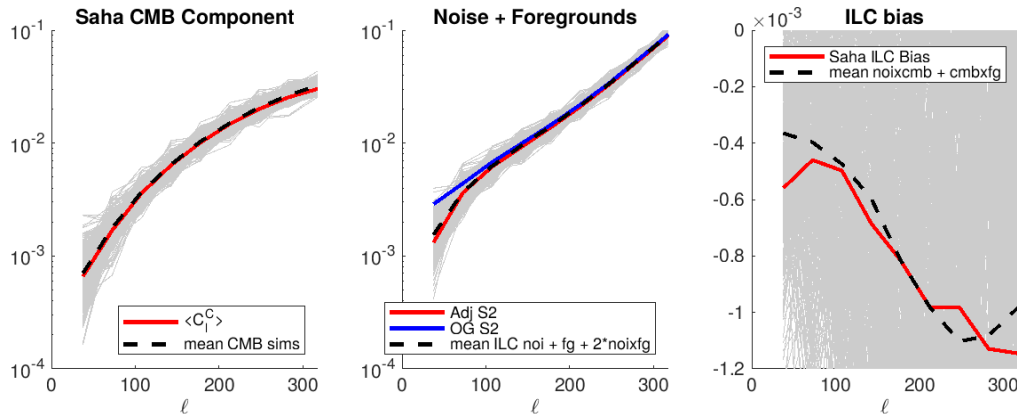
Figure 4.9: Left: $\langle \hat{D}^c_\ell \rangle$ compared with ILC CMB only realizations. Center: adjusted $S2$ in red compared with the original noise and foreground Saha term in blue, with the ILC dust, noise, and twice dust $\times$ noise realizations. Right: Saha ILC bias in red compared with ILC dust $\times$ CMB realizations, and CMB $\times$ noise.

and foreground term without the multiplicative term in blue and the adjusted noise and foreground term with the multiplicative factor in red. These two models are compared with the ILC noise power + ILC foreground power + $2 \cdot$ noise $\times$ foregrounds, where the average of these simulations is the dotted black line. The third panel shows the ILC bias compared with the average of dust $\times$ CMB + noise $\times$ CMB cross spectra. While the mean of these simulations does not completely agree with the analytic ILC bias term, we nonetheless see the general behavior of these lines agrees to close approximation.

As mentioned previously, the BK15 standard simulations contain foregrounds of only Gaussian dust and no synchrotron, and is the reason why we only consider ILC dust simulations for these comparisons. In later sections we will explore a Saha model allowing for synchrotron when we analyze altdust simulations, however we do not directly compare the terms of the Saha model to individual foreground, noise and CMB components of these simulations as all of the simulations are mixtures of all signals in the simulated sky.

Saha ILC Bias

In the simulations based model the bias inherent to the process of the ILC was addressed by examining the CMB $\times$ foregrounds and CMB $\times$ noise cross terms. In the Saha model this bias is addressed in the third term of equation (4.5), rewritten below:

$$bias = (1 - n_c)\frac{\langle \hat{D}_\ell^c \rangle}{N_{mode}} \tag{4.6}$$

Because we observe in more than one frequency band this bias is negative, and is more negative as the number of frequency channels grows.

Appendix E of [36] extensively derives the Saha ILC bias for the ideal case of foregrounds following a rigid frequency scaling, no detector noise, and the number of channels is $n_c \geq n_f + 1$, where n_f is the number of foreground components.

The case for $n_c \geq n_f + 1$ is especially true for our analysis, as we consider either $n_f = 1$, assuming foregrounds only composed of dust, or $n_f = 2$, assuming synchrotron is present as well. If we include an additional term for the CMB, this gives us a case of 3 independent components. The number of channels we have is 12, coming from the 12 frequency bands of the BK15 data set. It is often the case where the number of channels n_c is chosen to be equal to the full rank $n_f + 1$ of the covariance matrix to ensure the matrix is invertible, however we can see that it is important to consider the case of a singular covariance matrix arising from the fact that the number of channels $n_c = 12$ is greater than full rank, $n_f + 1 = 3$. Assuming a case where no detector noise is present is convenient for the sake of deriving a simplified form of the ILC bias, however the calculations can be generalized to include detector noise as well to give us the bias as written in equation (4.6).

We begin with the premise that the ILC power spectrum can be written of the form

$$\hat{C}_\ell^{Clean} = \frac{1}{\mathbf{e_0^T R}_\ell^{-1} \mathbf{e_0}} \tag{4.7}$$

where $\mathbf{R}_\ell$ is the covariance matrix described in section 3.4, and for singular matrix $\mathbf{R}_\ell$ can be rewritten

$$\hat{C}_\ell^{Clean} = \frac{1}{\mathbf{e_0^T R}_\ell^{\dagger} \mathbf{e_0}} \tag{4.8}$$

We now work through the steps as written in Appendix E of the Saha paper. In obtaining an equation for the ILC bias, we wish to explore further simplification of equation (4.8). To begin the calculation of the ILC bias under the ideal conditions described above, we need to first obtain an analytic expression for the covariance matrix $\mathbf{R}_\ell$. It is found that we can get the full covariance matrix through three successive rank-one updates to three separate matrices. We can then apply the Sherman Morrison formula to each of the three rank-one updates in order to obtain the analytic expression for our ideal foreground only ILC bias.

While until this point we have been writing the angular power spectrum in terms of D_ℓ, for the sake of remaining consistent with the equations in appendix E of [36] I will denote the power spectrum in C_ℓ until the end, when I will revert back to D_ℓ.

We recall that our full covariance matrix contains the auto and cross spectra between our frequency maps, thus we begin with a general equation for the auto power spectrum of the ith frequency channel:

$$\hat{C}_\ell^i = \hat{C}_\ell^c + 2\sum_{p=1}^{n_f} f_p^i \hat{C}_\ell^{cf(p)0} + \sum_{p,p'}^{n_f} f_p^i f_{p'}^i C_\ell^{(pp')0} \tag{4.9}$$

where $\hat{C}_\ell^c$ is the CMB auto power spectrum and f_p^i carries the frequency dependence of the pth foreground component for channel i. $\hat{C}_\ell^{cf(p)0}$ is the chance correlation between the CMB and the pth foreground component, and $C_\ell^{(pp')0}$ is the correlation between the two foreground

components p and p'. We can write the cross power spectrum between two frequency channels i and j as

$$\hat{C}_\ell^{i,j} = \hat{C}_\ell^c + \sum_{p=1}^{n_f} f_p^i \hat{C}_\ell^{cf(p)0} + \sum_{p=1}^{n_f} f_p^j \hat{C}_\ell^{cf(p)0} + F^{ij} \tag{4.10}$$

where

$$F^{ij} = \sum_{p,p'}^{n_f} f_p^i f_{p'}^j C_\ell^{(pp')0} \tag{4.11}$$

If we define the $\hat{\mathbf{f}}_\ell^{\mathbf{p0}}, \mathbf{e_0} \in \mathbb{R}_{n_c,1}$ vectors, where $\mathbb{R}_{n_c,n_f}$ is the set of real $n_c \times n_f$ matrices

$$\hat{\mathbf{f}}_\ell^{\mathbf{p0}} = \hat{C}_\ell^{cf(p)0} \begin{pmatrix} f_p^1 \\ f_p^2 \\ . \\ f_p^{n_c} \end{pmatrix}$$

$$\mathbf{e_0} = \begin{pmatrix} 1 \\ 1 \\ . \\ 1 \end{pmatrix} \tag{4.12}$$

We can write our full covariance matrix as

$$\mathbf{R}_\ell = \hat{C}_\ell^c \mathbf{e_0} \mathbf{e_0^T} + \left(\sum_{p=1}^{n_f} \hat{\mathbf{f}}_\ell^{\mathbf{p0}} \right) \mathbf{e_0^T} + \mathbf{e_0} \left(\sum_{p=1}^{n_f} \hat{\mathbf{f}}_\ell^{\mathbf{p0}} \right)^T + \mathbf{A_3} \tag{4.13}$$

where $\mathbf{A_3}$ is the covariance matrix between the foregrounds in each channel. We define the vector $\hat{\mathbf{f}}_\ell^{\mathbf{0}} \in \mathbb{R}_{n_c,1}$ to contain the cross of CMB and foreground components.

$$\hat{\mathbf{f}}_\ell^0 = \sum_{p=1}^{n_f} \hat{\mathbf{f}}_\ell^{p0} \tag{4.14}$$

We can rewrite our covariance matrix

$$\mathbf{R}_\ell = \hat{C}_\ell^c \mathbf{e_0} \mathbf{e_0^T} + \hat{\mathbf{f}}_\ell^0 \mathbf{e_0^T} + \mathbf{e_0} \hat{\mathbf{f}}_\ell^{0T} + \mathbf{A_3} \tag{4.15}$$

where

$$\mathbf{A_2} = \mathbf{e_0} \hat{\mathbf{f}}_\ell^{0T} + \mathbf{A_3}$$
$$\mathbf{A_1} = \hat{\mathbf{f}}_\ell^0 \mathbf{e_0^T} + \mathbf{A_2} \tag{4.16}$$

Now that we have this expression in terms of several successive rank-one updates, we need to obtain an analytic form of its inverse. Reference [37] works through several cases in which a general formula may be obtained of the inverse of matrix $\mathbf{M} = \mathbf{A} + \mathbf{b}\mathbf{c}^*$, where complex matrix $\mathbf{M}$ is the modification of the complex $m \times n$ matrix $\mathbf{A}$, and $\mathbf{b}$ and $\mathbf{c}$ are complex column vectors of size $m \times 1$ and $n \times 1$, respectively. Below I will detail the steps for which the Moore-Penrose General Inversion, or MPGI, may be calculated for our work as:

$$\mathbf{M}^\dagger = \mathbf{A}^\dagger + \mathbf{G} \tag{4.17}$$

where $\mathbf{G}$ is a matrix that contains only sums and products of the matrices $\mathbf{A}$, $\mathbf{A}^\dagger$, $\mathbf{b}$, $\mathbf{c}$ and their conjugate transposes.

Step 1

Here we consider the case where $\mathbf{b} \in \mathcal{C}(\mathbf{A})$, $\mathbf{c} \in \mathcal{C}(\mathbf{A}^*)$ and $\lambda \neq 0$, where $\mathcal{C}(\mathbf{A})$ is the column space for matrix $\mathbf{A}$ and $\lambda = 1 + \mathbf{c}^* \mathbf{A}^\dagger \mathbf{b}$. For this case we may write $\mathbf{M}^\dagger$ as

$$\mathbf{M}^{\dagger} = \mathbf{A}^{\dagger} - \lambda^{-1}\mathbf{de}^{*} \tag{4.18}$$

where

$$\mathbf{d} = \mathbf{A}^{\dagger}\mathbf{b}$$
$$\mathbf{e} = (\mathbf{A}^{\dagger})^{*}\mathbf{c} \tag{4.19}$$

From equation (4.15) we see we can write our covariance matrix

$$\mathbf{R}_{\ell} = \hat{C}_{\ell}^{c}\mathbf{e_0}\mathbf{e_0^T} + \mathbf{A_1} \tag{4.20}$$

where we note that $\mathbf{e_0} \in \mathcal{C}(\mathbf{A_1})$, and we wish to rewrite equation (4.8) in terms of $\mathbf{A_1}$. It can be shown that performing such a substitution will give us

$$\hat{C}_{\ell}^{Clean} = \hat{C}_{\ell}^{c} + \frac{1}{\mathbf{e_0^T}\mathbf{A_1^{\dagger}}\mathbf{e_0}} \tag{4.21}$$

Step 2

In this next step we will express the second term in equation (4.21) in terms of $\mathbf{A_2}$, where we recall $\mathbf{A_1} = \mathbf{f}_{\ell}^{0}\mathbf{e_0^T} + \mathbf{A_2}$, $\mathbf{f}_{\ell}^{0} \notin \mathcal{C}(\mathbf{A_2})$ and $\mathbf{e_0} \notin \mathcal{C}(\mathbf{A_2^T})$. Going back to our general notation, if we have our column vectors $\mathbf{b} \notin \mathcal{C}(\mathbf{A})$ and $\mathbf{c} \notin \mathcal{C}(\mathbf{A}^{*})$, then $\mathbf{M}^{\dagger}$ may be written as

$$\mathbf{M}^{\dagger} = \mathbf{A}^{\dagger} - \mathbf{ku}^{\dagger} - \mathbf{v}^{\dagger}\mathbf{h} + \lambda\mathbf{v}^{\dagger}\mathbf{u}^{\dagger} \tag{4.22}$$

where

$$\mathbf{k} = \mathbf{A}^\dagger \mathbf{b}$$

$$\mathbf{u} = (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)\mathbf{b}$$

$$\mathbf{v} = \mathbf{c}^*(\mathbf{I} - \mathbf{A}^\dagger \mathbf{A}) \tag{4.23}$$

$$\mathbf{h} = \mathbf{c}^*\mathbf{A}^\dagger$$

Then we can write

$$\mathbf{A}_1^\dagger = \mathbf{A}_2^\dagger - \mathbf{k}\mathbf{u}^\dagger - \mathbf{v}^\dagger\mathbf{h} + \lambda\mathbf{v}^\dagger\mathbf{u}^\dagger \tag{4.24}$$

where

$$\mathbf{k} = \mathbf{A}_2^\dagger \mathbf{f}_\ell^0$$

$$\mathbf{h} = \mathbf{e}_o^{\mathbf{T}}\mathbf{A}_2^\dagger$$

$$\mathbf{u}^\dagger = \frac{\mathbf{f}_\ell^{0\mathbf{T}}(\mathbf{I} - \mathbf{A}_2\mathbf{A}_2^\dagger)}{\mathbf{f}_\ell^{0\mathbf{T}}(\mathbf{I} - \mathbf{A}_2\mathbf{A}_2^\dagger)\mathbf{f}_\ell^0} \tag{4.25}$$

$$\mathbf{v}^\dagger = \frac{(\mathbf{I} - \mathbf{A}_2^\dagger\mathbf{A}_2)\mathbf{e}_0}{\mathbf{e}_0^{\mathbf{T}}(\mathbf{I} - \mathbf{A}_2^\dagger\mathbf{A}_2)\mathbf{e}_0}$$

and we skip some steps to obtain

$$\frac{1}{\mathbf{e}_0^{\mathbf{T}}\mathbf{A}_1^\dagger\mathbf{e}_0} = \frac{\mathbf{f}_\ell^{0\mathbf{T}}(\mathbf{I} - \mathbf{A}_2\mathbf{A}_2^\dagger)\mathbf{f}_\ell^0}{\mathbf{f}_\ell^{0\mathbf{T}}(\mathbf{I} - \mathbf{A}_2\mathbf{A}_2^\dagger)\mathbf{e}_0} \tag{4.26}$$

Step 3

Next, we consider $\mathbf{b} \notin \mathcal{C}(\mathbf{A})$ and $\mathbf{c} \in \mathcal{C}(\mathbf{A}^*)$, then

$$\mathbf{M}\mathbf{M}^\dagger = \mathbf{A}\mathbf{A}^\dagger - \eta^{-1}\mathbf{e}\mathbf{e}^{\mathbf{T}} + \eta^{-1}\nu^{-1}\mathbf{q}\mathbf{q}^{\mathbf{T}} \tag{4.27}$$

where

$$\mathbf{e} = (\mathbf{A}^\dagger)^* \mathbf{c}$$

$$\eta = \mathbf{e}^* \mathbf{e}$$

$$\nu = \lambda\lambda^* + \eta\phi \qquad (4.28)$$

$$\phi = \mathbf{f}^\mathbf{T} \mathbf{f}$$

$$\mathbf{f} = (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)\mathbf{b}$$

In this case we are going to rewrite equation (4.26) in terms of $\mathbf{A_3}$, and we do this by first expressing $\mathbf{A_2}\mathbf{A_2}^\dagger$ in terms of $\mathbf{A_3}\mathbf{A_3}^\dagger$ recalling that $\mathbf{A_2} = \mathbf{A_3} + \mathbf{e_0}\mathbf{f_0}^\mathbf{T}$ and $\mathbf{f}_\ell^\mathbf{0} \in \mathcal{C}(\mathbf{A_3}^*), \mathbf{e_0} \notin \mathcal{C}(\mathbf{A_3})$. We can then write

$$\mathbf{A_2}\mathbf{A_2}^\dagger = \mathbf{A_3}\mathbf{A_3}^\dagger - \eta^{-1}\mathbf{e}\mathbf{e}^\mathbf{T} + \eta^{-1}\nu^{-1}\mathbf{q}\mathbf{q}^\mathbf{T} \qquad (4.29)$$

where

$$\mathbf{e} = \mathbf{A_3}^\dagger \mathbf{f}_\ell^\mathbf{0}$$

$$\mathbf{q} = \lambda\mathbf{e} + \eta\mathbf{f} \qquad (4.30)$$

$$\mathbf{f} = (\mathbf{I} - \mathbf{A_3}\mathbf{A_3}^\dagger)\mathbf{e_0}$$

Again we have everything we need to proceed and skip some steps to obtain the second term in the numerator of equation (4.26)

$$\mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_2}\mathbf{A_2}^\dagger \mathbf{f}_\ell^\mathbf{0} = \mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3}\mathbf{A_3}^\dagger \mathbf{f}_\ell^\mathbf{0} - \eta^{-1}(\mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3}^\dagger \mathbf{f}_\ell^\mathbf{0})^2 + \eta^{-1}\nu^{-1}(\lambda^2(\mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3}^\dagger \mathbf{f}_\ell^\mathbf{0})^2 +$$
$$\eta^2(\mathbf{f}_\ell^{\mathbf{0T}}(\mathbf{I} - \mathbf{A_3}\mathbf{A_3}^\dagger)\mathbf{e_0})^2 + 2\lambda\eta(\mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3}^\dagger \mathbf{f}_\ell^\mathbf{0} \mathbf{f}_\ell^{\mathbf{0T}}(\mathbf{I} - \mathbf{A_3}\mathbf{A_3}^\dagger)\mathbf{e_0})) \qquad (4.31)$$

Conveniently several terms drop from this equation due to the fact that $\mathbf{f}_\ell^\mathbf{0} \in \mathcal{C}(\mathbf{A_3})$, so $(\mathbf{A_3}\mathbf{A_3}^\dagger)\mathbf{f}_\ell^\mathbf{0} = \mathbf{f}_\ell^\mathbf{0}$. Using the definitions for ν and ϕ given in equation (4.28) and define the

variables X and Y

$$X = \mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3} \mathbf{A_3}^\dagger \mathbf{f}_\ell^{\mathbf{0}}$$
$$Y = \mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3}^\dagger \mathbf{f}_\ell^{\mathbf{0}} \tag{4.32}$$

we can write

$$\mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_2} \mathbf{A_2}^\dagger \mathbf{f}_\ell^{\mathbf{0}} = X + \frac{Y^2 \phi}{\nu} \tag{4.33}$$

We can then rewrite the numerator of equation (4.26)

$$\mathbf{f}_\ell^{\mathbf{0T}} (\mathbf{I} - \mathbf{A_2} \mathbf{A_2}^\dagger) \mathbf{f}_\ell^{\mathbf{0}} = \frac{-Y^2 \phi}{\nu} \tag{4.34}$$

and in a similar fashion, the denominator

$$\mathbf{f}_\ell^{\mathbf{0T}} (\mathbf{I} - \mathbf{A_2} \mathbf{A_2}^\dagger) \mathbf{e_0} = \frac{Y \phi}{\nu} \tag{4.35}$$

We see once we input equations (4.34) and (4.35) into equation (4.26), we obtain

$$\frac{1}{\mathbf{e_0^T} \mathbf{A_1}^\dagger \mathbf{e_0}} = -Y = -\mathbf{f}_\ell^{\mathbf{0T}} \mathbf{A_3}^\dagger \mathbf{f}_\ell^{\mathbf{0}} \tag{4.36}$$

Referring back to equation (4.21), the bias to the ILC is the quantity we have just computed in equation (4.36).

Obtaining Full Bias Expression

Now that we have our bias in terms of the foreground vectors and the updates to the covariance matrix defined in equations (4.14), (4.15), and 4.16 we can rewrite in terms of the power spectra. We begin with writing the foreground covariance matrix in terms of its

elements

$$A_3^{ij} = \sum_{pp'}^{n_f} f_p^i f_{p'}^j C_\ell^{(pp')0} \tag{4.37}$$

and recall the definition of the vector containing chance correlations between the CMB and foregrounds

$$\mathbf{f}_\ell^{0i} = \sum_p^{n_f} \hat{C}_\ell^{(cp)0} f_p^i \tag{4.38}$$

The magnitude of the ILC bias as given in equation (4.36) becomes

$$\langle \mathbf{f}_\ell^{0\mathbf{T}} \mathbf{A}_3^\dagger \mathbf{f}_\ell^0 \rangle = \sum_{ij} A_3^{\dagger ij} \left\langle \sum_{pp'} \hat{C}_\ell^{(cp)0} \hat{C}_\ell^{(cp')0} \right\rangle f_p^i f_{p'}^j \tag{4.39}$$

From equation 1.15 we can infer that $\hat{C}_\ell^{(cp)0} = \sum_m \frac{a_{\ell,m}^c a_{\ell,m}^{p0*}}{2\ell+1}$, and we know that the CMB anisotropies are isotropic so $\langle a_{\ell,m} a_{\ell',m'} \rangle = C_\ell^c \delta_{\ell,\ell'} \delta_{m,m'}$, we get

$$\left\langle \sum_{pp'} \hat{C}_\ell^{(cp)0} \hat{C}_\ell^{(cp')0} \right\rangle = \frac{C_\ell^c}{2\ell+1} \sum_{pp'} C_\ell^{(pp')0} \tag{4.40}$$

Substituting equations (4.37) and (4.40) in for equation (4.39) gives us

$$\langle \mathbf{f}_\ell^{0\mathbf{T}} \mathbf{A}_3^\dagger \mathbf{f}_\ell^0 \rangle = \frac{C_\ell^c}{2\ell+1} \sum_{ij} A_3^{\dagger ij} A_3^{ij} \tag{4.41}$$

We note that $\sum_{ij} A_3^{\dagger ij} A_3^{ij} = tr(\mathbf{A}_3^\dagger \mathbf{A}_3) = rank(\mathbf{A}_3) = n_f$, so

$$\langle \mathbf{f}_\ell^{0\mathbf{T}} \mathbf{A}_3^\dagger \mathbf{f}_\ell^0 \rangle = nf \frac{\langle \hat{C}_\ell^c \rangle}{2\ell+1} \tag{4.42}$$

Now we can rewrite equation 4.21 with this substitution and in terms of D_ℓ in place of

C_ℓ as

$$\langle \hat{D}_\ell^{Clean} \rangle = \langle \hat{D}_\ell^c \rangle - nf \frac{\langle \hat{D}_\ell^c \rangle}{2\ell + 1} \qquad (4.43)$$

We recall that this equation for the ILC power spectrum has been derived under the conditions that detector noise is not present, and that foregrounds follow a rigid frequency scaling. Once we consider a case where foregrounds have a more complicated relationship with frequency (i.e. dust follows a power law in frequency and synchrotron follows a separate frequency power law) and detector noise is present, a positive bias appears along with the negative bias derived above and we get equation 4.4.

4.2 Hamimeche-Lewis Likelihood Approximation

The likelihood used for the ML search analyses is the Hamimeche-Lewis (H-L) log likelihood approximation [38], and is the same likelihood calculation used by the standard BICEP/Keck analysis. This is a useful method for the data we take because it was specifically developed for partial sky observations and can obtain correct results at low ℓ, and is defined

$$-2log\mathcal{L}(\{\mathbf{C}_\ell\}|\{\hat{\mathbf{C}}_\ell\}) = \sum_{\ell\ell'}[\mathbf{X}_g]_\ell^T[\mathbf{M}_f^{-1}]_{\ell\ell'}[\mathbf{X}_g]_{\ell'} \qquad (4.44)$$

Here $\mathbf{C}_\ell$ are the bandpowers of the model being tested, $\hat{\mathbf{C}}_\ell$ are the data bandpowers, $\mathbf{M}_f$ is the fiducial bandpower covariance matrix, and $\mathbf{X}_g$ is the expression of the bandpowers:

$$[\mathbf{X}_g]_\ell = \text{vecp}(\mathbf{C}_{f\ell}^{1/2}\mathbf{g}[\mathbf{C}_\ell^{-1/2}\hat{\mathbf{C}}_\ell\mathbf{C}_\ell^{-1/2}]\mathbf{C}_{f\ell}^{1/2}) \qquad (4.45)$$

with the nonlinear function $\mathbf{g}$:

$$g(x) \equiv \text{sign}(x-1)\sqrt{2(x - \ln(x) - 1)} \qquad (4.46)$$

where $\mathbf{C}_{f\ell}$ are the fiducial model bandpowers, or the mean of simulations, vecp() is a vector function giving the vector of unique elements in a symmetric matrix, and by definition the application of $\mathbf{g}$ to a positive definite matrix is equivalent to applying g to its eigen-values. With this likelihood function we are then able to test the ILC models described in sections 4.1.1 and 4.1.3.

In terms of this specific analysis, the inputs required by the H-L likelihood to be calculated at any point in parameter space are:

- The fiducial bandpower covariance matrix $\mathbf{M}_f$ of the 9 science ℓ bins that we have been investigating. We zero out covariance between ℓ bins not adjacent to the diagonal because although there is no correlation expected to appear between non-adjacent bins, in practice there are some chance correlations that do appear.

- Fiducial bandpowers, $\mathbf{C}_{f\ell}$, which are the average of the ILC realizations, or the black line in figure 4.1.

- Data, or $\hat{\mathbf{C}}_\ell$, which for this analysis are individual ILC realizations, or the grey lines in figure 4.7.

- ILC model, or $\mathbf{C}_\ell$.

4.3   Sims based Likelihood/Posterior Calculation

We created ILC spectra for a range of A_d, β_d, and r values using standard Gaussian dust simulations and want to recover the true values of these parameters. By construction, the ILC is insensitive to A_d and β_d because it removes the foregrounds, and once the maps go

through the ILC process it is hard to distinguish between the effects of changing A_d versus changing β_d. In an effort to find the true values, we need to calculate a posterior, which places reasonable constraints on the likelihood and tells us the chance that a model that we fit to a given measurement of data is true. The posterior can be calculated by multiplying the likelihood by a prior distribution, which as its name suggests is a distribution based upon prior knowledge of the given variable. For the simulations based likelihood our posterior is:

$$\text{posterior} = P(\beta_d)\mathcal{L} \tag{4.47}$$

The prior on β_d is the Gaussian prior used by BICEP/Keck shown in in equation (4.48), where $\sigma_{\beta_d} = 0.11$ and $\mu_{\beta_d} = 1.6$. We do not have a good model for dust amplitude in our particular patch of sky so do not include a prior on A_d in the posterior calculation.

$$P(\beta_d) = \frac{1}{\sigma_{\beta_d}\sqrt{2\pi}} \exp\left(-\frac{(\beta_d - \mu_{\beta_d})^2}{(2\sigma_{\beta_d}^2)}\right) \tag{4.48}$$

Calculating the posterior at every point in the 3-dimensional parameter space of A_d, β_d and r is a very time consuming process, so rather than performing a grid calculation to get a full posterior shape we can do a Metropolis-Hastings Monte Carlo Markov Chain (MH-MCMC). Figure 4.10 plots the MCMC posterior curves for a few ILC realizations with the nominal values of $A_d = 3.75$ and $\beta_d = 1.6$.

The MCMC plots were created in Python using the package getdist, which is an MCMC engine used to perform cosmological parameter estimation, and in this case was used for plotting posterior MCMC chains calculated in MATLAB. While the full posterior is not necessary in an ML search, seeing the full shape is useful.

We see in the first column of plots, going from top to bottom, the marginalized β_d posterior curved where we have integrated over all values of A_d and r, in the center the two dimensional posterior for A_d and β_d where we have integrated over r, and on the bottom the

two dimensional posterior for r and β_d where we have integrated over A_d. In these plots we see the curves centered pretty close to $\beta_d = 1.6$, which we would expect because the value of β_d in the given data is 1.6, and because we center our very tight β_d prior about $\beta_d = 1.6$. We can also see that the spread of the distributions is very narrow in β_d, which is due to the prior we've provided for β_d being so narrow. While this distribution is calculated in parameter space for three variables (A_d, β_d and r), because the prior on β_d is so tight this posterior realistically only has 2 independent variables (A_d and r). In the center plot we see a wide spread in A_d, which is due to the loose constraint we have placed on this parameter.

In the second column of figure 4.10 we see on top the marginalized A_d posterior where we have integrated over β_d and r space. We note that the realization shown as the black line has a smaller local maxima, which is because the kernel used to smooth out this curve was chosen to zero the posterior for $A_d = 10$, as we have reasonable expectation that the true value of A_d lies within the range of 0 to 10. The bottom plot in this column shows the two dimensional posterior for A_d and r where β_d has been integrated out. We see that the posterior is narrower in r than in the A_d direction.

Finally, the lone plot in the third column shows the marginalized r posterior curve where we have integrated over A_d and β_d. We see that these curves peak a bit below zero. If we were to plot all 499 realizations, we would expect that their peaks would spread to both positive and negative r values, and if our results are unbiased, the distribution of negative and positive peaks should each be roughly half of the realizations.

As mentioned above, while it is useful to look at plots of the full posterior curves as a diagnostic, the information that we are most interested in is the distribution of the maxima of the posteriors for all realizations.
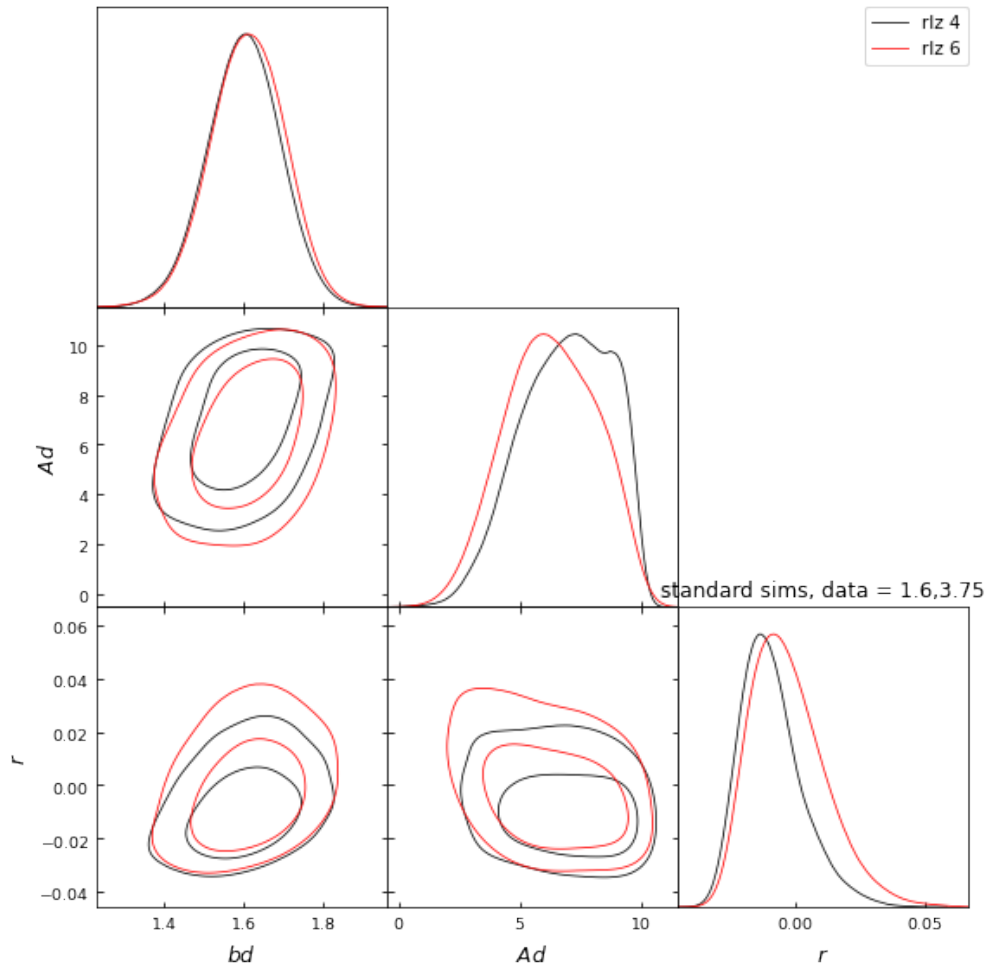
Figure 4.10: MCMC posterior triangle plot for simulation based ILC model. Top left: marginalized β_d posterior, where A_d and r have been integrated over. Middle left: 2D A_d and β_d posterior where r has been integrated out. Bottom left: 2D r and β_d posterior where A_d has been integrated out. Center top: marginalized A_d posterior, where β_d and r have been integrated out. Center bottom: 2D r and A_d posterior, where β_d has been integrated out. Right: marginalized r posterior, where A_d and β_d have been integrated out.

4.4   Sims ILC model Maximum Likelihood Search

We may conduct a maximum likelihood search using the posterior in place of likelihood. Because the posterior is a prior multiplied by the likelihood, finding the maximum of a likelihood is the same as searching for the maximum of a posterior. As mentioned before, we only have an explicit prior for β_d, however we have a reasonable expectation of where A_d and r should exist, therefore we only allow our search to explore between $A_d = [0,10]$, and $r = [-0.05,0.15]$, with an additional constraint for β_d to be within the values of 1 and 3 by forcing the posterior to infinity at these points. [3] The BICEP/Keck ML search algorithm makes use of the MINUIT library developed by CERN. We also note that our simulation based model is best defined within this parameter space. There are some points in the allowed parameter space where the calculated likelihood is imaginary, and we don't allow the ML search to explore these regions.

To test the robustness of our ILC model, we use standard simulations created with varying A_d and β_d values (the grey lines in figure 4.7). We are also able to use ILC power spectra composed from altdust simulations, where alternative models are used to describe foregrounds. The four altdust models studied in this analysis are described in Table 4.1. Some of these simulations predict higher foreground levels than we observe in our patch of sky, however they prove useful as a test of the performace of our ILC analysis.

In three of the four altdust models in table 4.1 we see in the Python sky models (PySM) a series of numbers and letters, for example in the 1682×1683 altdust model an alternative name is PySM model a2d4f1s3. The letters in the sequence indicate the different foreground components: AME (a), dust (d), free-free (f), and synchrotron (s). The numbers refer to the models of each foreground component described in the PySM paper [39]. The a1 and f1 models are unpolarized and are thus irrelevant to our analysis. The a2 model has polarized AME, however because our analysis does not include this component we may expect to see

---

[3]While we talk about Maximum Likelihood Searches, the way the code actually works is to find find the minimum of the negative log likelihood, as this is computationally easier to do

Model	Description
1682x1683	PySM model a2d4f1s3. Two-component dust and curvature in synchrotron spectral index. Polarized AME.
1688x1689	Flauger–Hensley MHD simulations that are meant to provide a consistent simulation of dust and synchrotron.
1691x1692	PySM model a1d1f1s1. Most basic PySM model including dust, synchrotron, AME, and free-free.
1693x1694	PySM model a2d7f1s3. Hensley–Draine dust with more complicated SED. Curvature in synchrotron spectral index. Polarized AME.

Table 4.1: Summary of the four types of altdust model simulations analyzed with the ILC. Models include varying levels of foreground and synchrotron signal. Some models include contribution of anomalous microwave emission (AME), which is an unexpected component of micrwave emission observed in the 10-60 GHz range and as can be inferred from its name, the source of AME is an open question.

some bias from these simulations.

The d1 model uses dust templates from the Planck COMMANDER code [40] of the 545 GHz intensity and 353 GHz polarization, which is highly dominated by dust. They find a best COMMANDER fit assuming frequency scaling as a single component, and the intensity and polarization share the same index. β_d and T_d vary spatially so the model implements some dust pattern decorrelation, which in practice is quite small. The d4 model is a more generalized version of d1 where it assumes two dust components each with its own temperature, using templates from Ref. [41]. The d7 model is a separate physical model of dust which does not conform to the modified blackbody SED and is described in Ref. [42].

The s1 model simulates the spectral index of synchrotron as a power law in every direction, and uses this to model the synchrotron polarization as a scaling of the 9 year WMAP Q and U maps, assuming the index is the same temperature and polarization, knowing this is a simplification of the true sky. The s3 model takes the s1 model and includes a curvature on the spectral index.

All Planck and WMAP maps are noise dominated at smaller angular scales at higher

Galactic altitude. The PySM model smooths the maps at these scales and inputs realizations with Gaussian foreground structure as described in chapter 3 of Ref. [39].

The second altdust model listed, 1688×1689, is based on Galactic magnetic field simulations [43] which naturally produces correlated dust and synchrotron, and gives a higher level of synchrotron than is favored by the BICEP/Keck data. These simulations are not based on external data so have no issue with noise level, and are quite different from the standard Gaussian dust simulations in that the structure of foregrounds from this model are non-Gaussian for all ℓ.

The standard BK15 analysis compares the simulations to an 8 parameter model of lensed ΛCDM+r+dust+synchrotron+noise, exploring the parameter space and conducting a maximum likelihood search using COSMOMC which implements a Markov chain Monte Carlo. We compare our 3 parameter maximum posterior search to the BICEP/Keck multi-spectrum ML search with the modification to only allow for the three free parameters, A_d, β_d, and r. Results are shown for the three parameter BK15 ML search in figure 4.14.

Table 4.2 compares the statistics of this search with those of the simulation based ILC maximum likelihood search using standard simulations with nominal values of $A_d = 3.75$ and $\beta_d = 1.6$. Note that while the means of the A_d and β_d maximum likelihood distributions are directly listed, it is $\bar{r}/\sigma_r$ which is shown for the r maximum likelihood distribution, as this is the standard of comparison for the BICEP/Keck analyses.

It is important to note that the effects that A_d and β_d have on the ILC power spectrum are less than their effects on the frequency simulations before combination, so we do not have as much constraining power on A_d as the standard BICEP/Keck maximum likelihood search and place more focus on marginalizing r. We can see evidence of this looking at the second column of table 4.2 showing the A_d results, where the ILC results give $\sigma_{A_d} = 1.79$ and the BICEP/Keck results give $\sigma_{A_d} = 0.68$. We see similar results for the β_d best fit between the ILC and BK analyses, which is due to the fact that the same tight β_d prior is

BICEP/Keck and Sims ILC Max Likelihood Search Results			
Model	A_d $(\bar{A}_d,\ \sigma_{A_d})$	β_d $(\bar{\beta}_d,\ \sigma_{\beta_d})$	r $(\sigma_r,\ \bar{r}/\sigma_r)$
BICEP/Keck	3.87, 0.68	1.6, 0.05	0.02, $+0.1\sigma$
Sims ILC	3.78, 1.79	1.6, 0.01	0.015, $+0.05\sigma$

Table 4.2: Results from standard BICEP/Keck 3 free parameter maximum likelihood search performed on standard simulations compared to the simulations based ILC maximum likelihood search. The ILC analysis proved to have a tighter constraint on σ_r, while being less sensitive to the dust parameters A_d and β_d as compared to the BK15 search.

used used both calculations. The simulations based ILC method finds less biased results for r and a smaller $\sigma_r = 0.015$ compared to the three free parameter standard BK ML search which finds $\sigma_r = 0.02$.

The plots in figures 4.11 - 4.13 show the simulations based ILC ML search results for ILC data composed with simulations having $\beta_d = 1.2, 1.6, 2$ and keeping fixed $A_d = 3.75$. Some interesting behavior arises as β_d increases from 1.2 to 2. We see the bias on r, or μ_r, decreasing from positive to negative. A source of this bias could be that as the value of β_d changes, the spread of ML values truncates at the edges in the A_d direction as evidenced in the scatter plots between A_d and r. Another source of this bias, especially at the extremes of $\beta_d = 1.2$ and $\beta_d = 2$, may be that the prior is strongly pulling β_d away from these values.

We can check the goodness of fit by looking at the distribution of $-2log(\mathcal{L})$ values. We expect the mean of this distribution to be approximate to the degrees-of-freedom, which is the number of data points (in our case 9, as we look at 9 ℓ bins) minus the number of model parameters (for us this is 2, A_d and r). As mentioned previously, because there is such a strong prior on β_d, this cannot be considered a true free parameter, and we expect the degrees of freedom to be 7. This is confirmed by calculating χ^2 at the values in parameter space given by the ML search, and can be seen in the top right panel in the ML search figures.

Figures 4.15 through 4.18 shows the ML search results for the altdust simulations. These results are less straightforward than those of the standard simulations, as some of the altdust

Figure 4.11: Simulations based ILC maximum likelihood search results for standard ILC input simulations composed with $\beta_d = 1.2$ and $A_d = 3.75$. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.
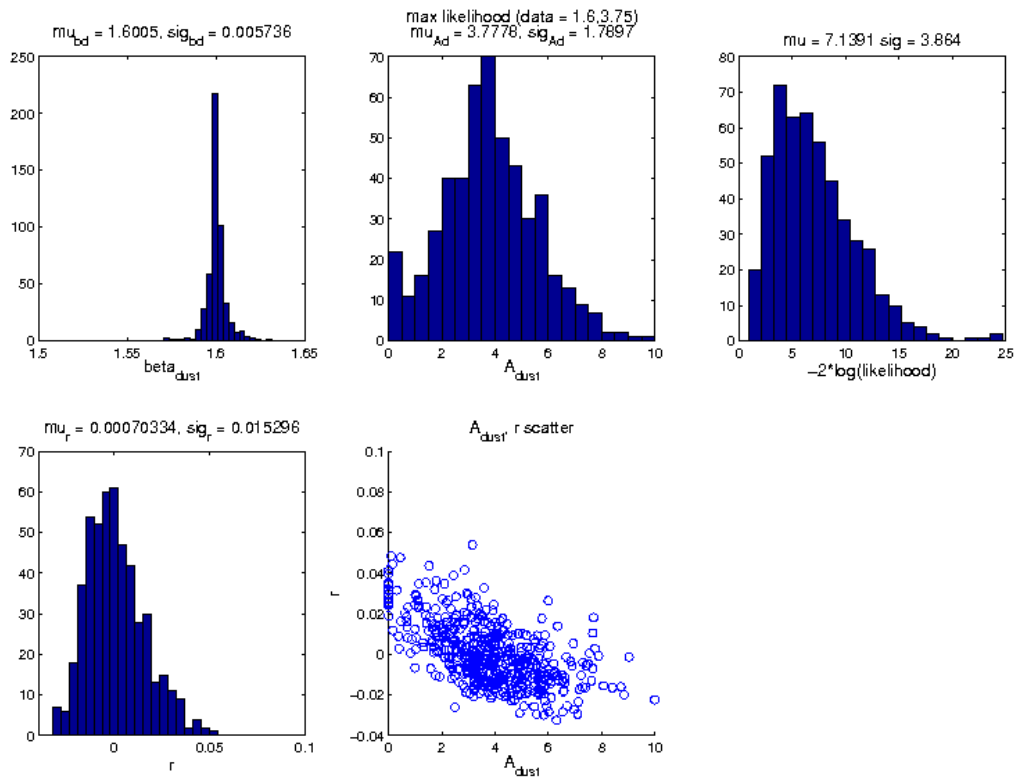
Figure 4.12: Simulations based ILC maximum likelihood search results for standard ILC input simulations composed with $\beta_d = 1.6$ and $A_d = 3.75$. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.
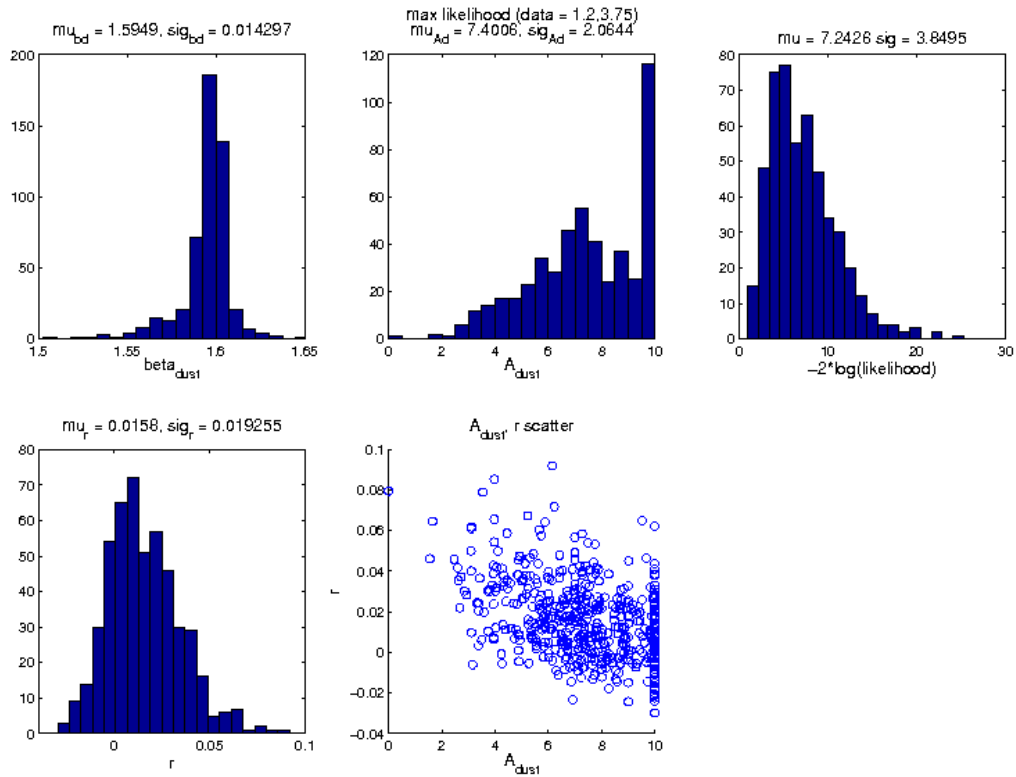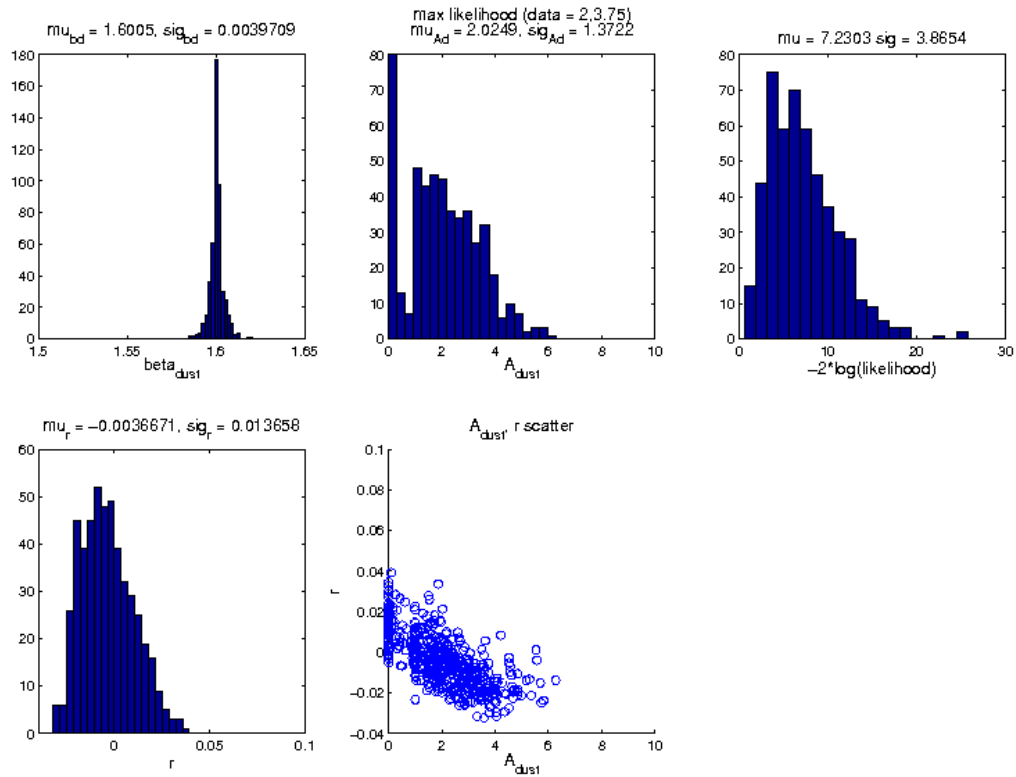
Figure 4.13: Simulations based ILC maximum likelihood search results for standard ILC input simulations composed with $\beta_d = 2$ and $A_d = 3.75$. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.
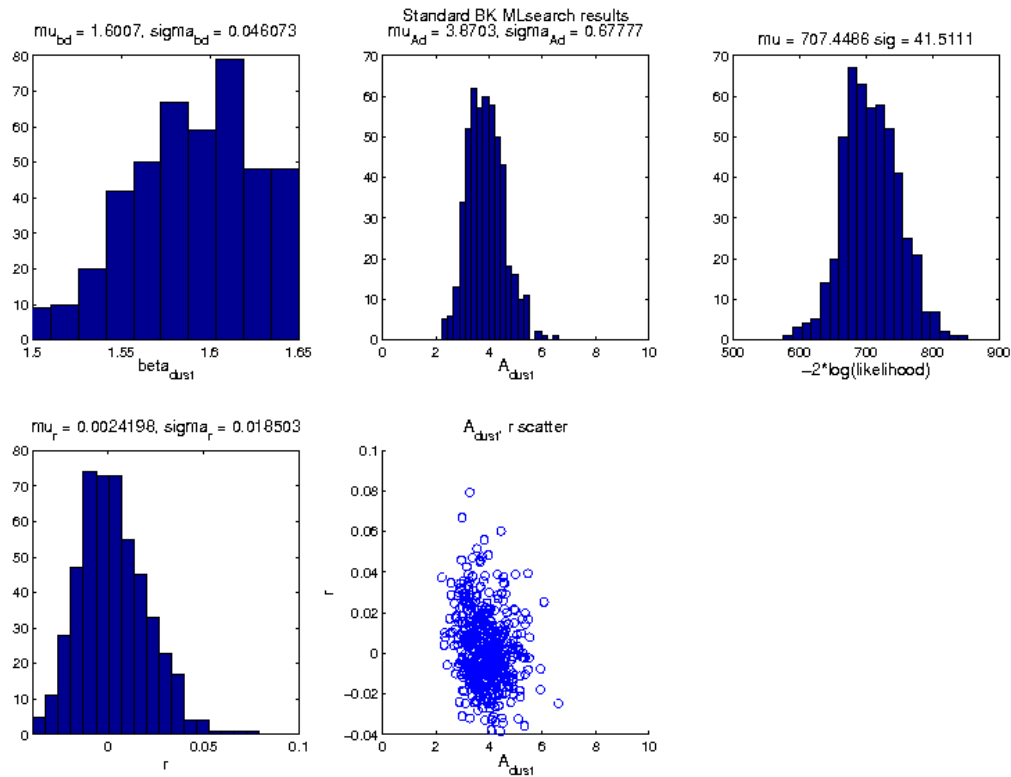
Figure 4.14: BICEP/Keck three free parameter maximum likelihood search results. Standard simulations were run through the standard BK analysis, however only the A_d, β_d, and r parameters were free while the remaining 5 were kept fixed. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.

simulations include polarized AME, and all models include synchrotron in their foregrounds as well.

While these simulations have varying models for their foreground compositions that are very different from the simple Gaussian dust and no synchrotron foregrounds of the standard simulations, our simulations based ILC model still has foreground and ILC biases calculated from the standard simulations. This is due to the fact that for the altdust models we only have the full sky simulations, and do not have the conveniently separated foreground only simulations like we do for our standard simulations. Therefore we keep in mind with these altdust ML results that while we know that the simulations contain synchrotron and other components, in some cases modeled in quite a complicated fashion, our simulation based ILC model only has the capability of assuming that foreground levels in the maps are of the same level as the Gaussian dust simulations.

Another important note is the ML search done on altdust simulations had noise, foreground, and ILC biases calculated with a lowess interpolation surface fit, rather than the 2D linear surface fit described for the standard simulations. The lowess fit is almost flat going from $A_d = 10$ to higher A_d, however the linear surface fit tended to increase very significantly for high A_d values. For these simulations, A_d was allowed to range from 0 to 50 and r was again limited to the range [-0.05,0.15].

Figure 4.15 shows the results for the 1682×1683 model, where there is present polarized AME, two-component dust, and a curvature in the synchrotron spectral index. In the top left panel we see that we find unbiased results for β_d with quite a narrow distribution, likely owing to the narrow constraint we place on this parameter. We find a best fit in the center top plot of $A_d = 28.15$. Our results for r are quite biased in the bottom left plot with $\mu_r = -0.75\sigma$, and the distribution appears to be sharply cut off at the lower end. If we look at the scatter plot in the center bottom, we have confirming evidence that the distribution of points is cut off at the lower end of our r range, which is where we place a boundary on

Figure 4.15: Simulations based ILC maximum likelihood search results for ILC input simulations composed from 1682×1683 altdust simulations. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.

our ML search. If we widened the ML search range in r it is likely that we would find more biased results in r. We find from our χ^2 plot in the top right corner shows a mean value of 8.67 where ideally this would be 7, telling us, unsurprisingly, that our model is not a very good fit to these simulations.

Figure 4.16 shows the results for the altdust model 1688×1689, which has correlated dust and synchrotron. These results show an unbiased β_d distribution due to the tight constraint placed on β_d. The A_d distribution gives a best fit value of $A_d = 3.13$, similar to the Gaussian dust simulations $A_d = 3.75$, however we see quite the pileup at $A_d = 0$, which is a boundary we placed on our ML search. This suggests that the model finds likely
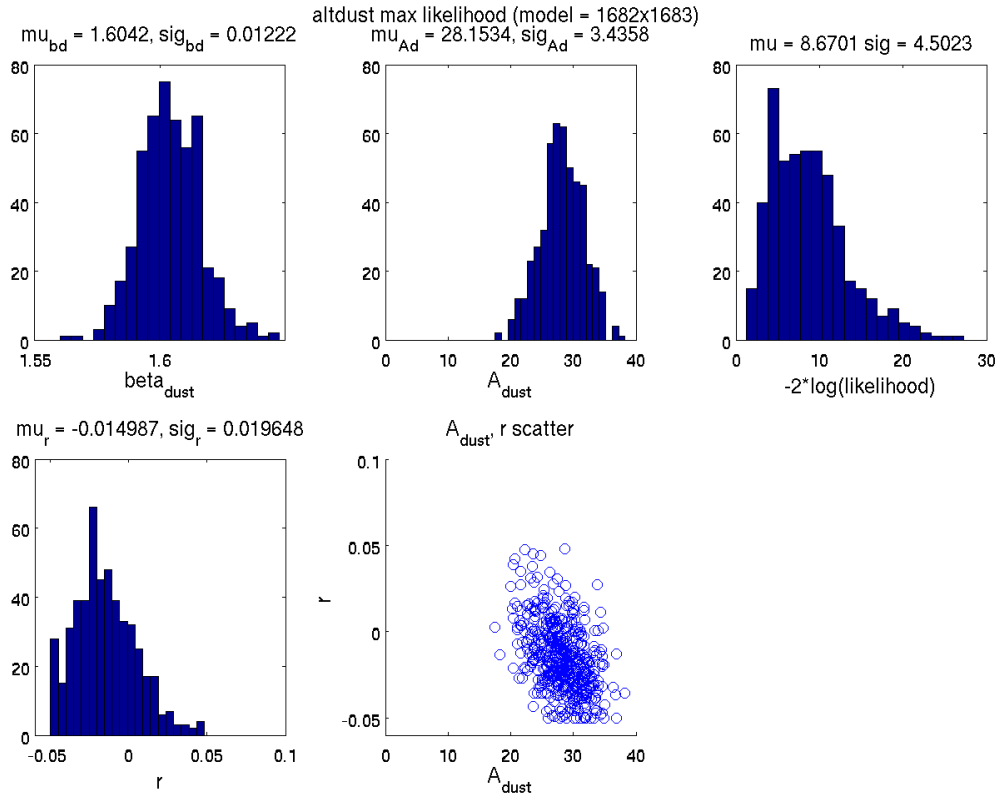
Figure 4.16: Simulations based ILC maximum likelihood search results for ILC input simulations composed from 1688×1689 altdust simulations. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.

values for the nonphysical region of negative A_d. The r results are quite biased as well, with $\mu_r = 0.4\sigma$, and our χ^2 distribution gives a mean value of 7. The results here look better than those from the 1682×1683, which may not be surprising because the 1688×1689 simulations do not have polarized AME. Still we obtain worse results than the standard simulations, as expected because these simulations more foreground components that are not accounted for in the simulations ILC model.

Figure 4.17 shows the results for the 1691×1692 simulations, which include synchrotron and a dust model where β_d varies spatially. These results give a best fit $A_d = 13.59$ and an unbiased r distribution with narrow $\sigma_r = 0.018$, though the χ^2 distribution for this model
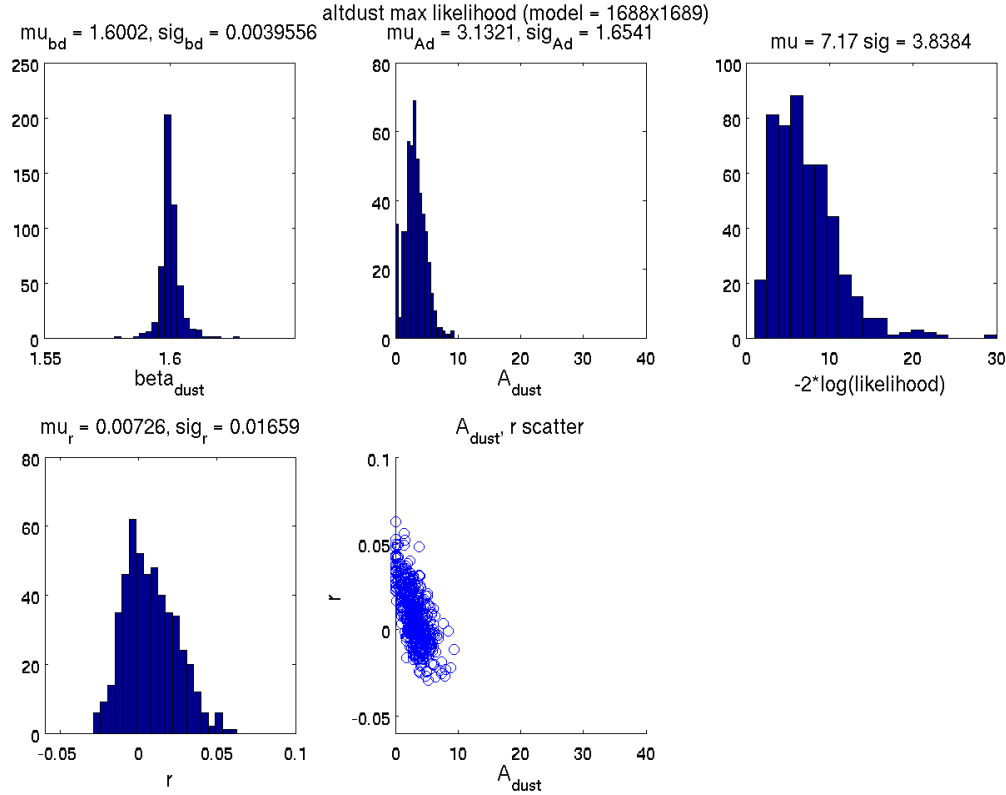
Figure 4.17: Simulations based ILC maximum likelihood search results for ILC input simulations composed from 1691×1692 altdust simulations. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.

is a bit high at 7.69. The β_d distribution for this model is wider than the previous ones we have seen although all ML searches have included the same β_d prior. It's possible the analysis had a harder time constraining β_d for this model because it assumes β_d is the same value across the area of sky contained in the simulations, while the altdust model assumes a varying value for β_d.

Figure 4.18 shows the results for altdust model 1693×1694 which are modeled with the complicated Hensley SED. While the model for the foregrounds in these simulations is quite different than the standard simulations which contain only Gaussian dust, we appear to find unbiased results for r. We find a best fit $A_d = 18.38$ and a β_d distribution centered about
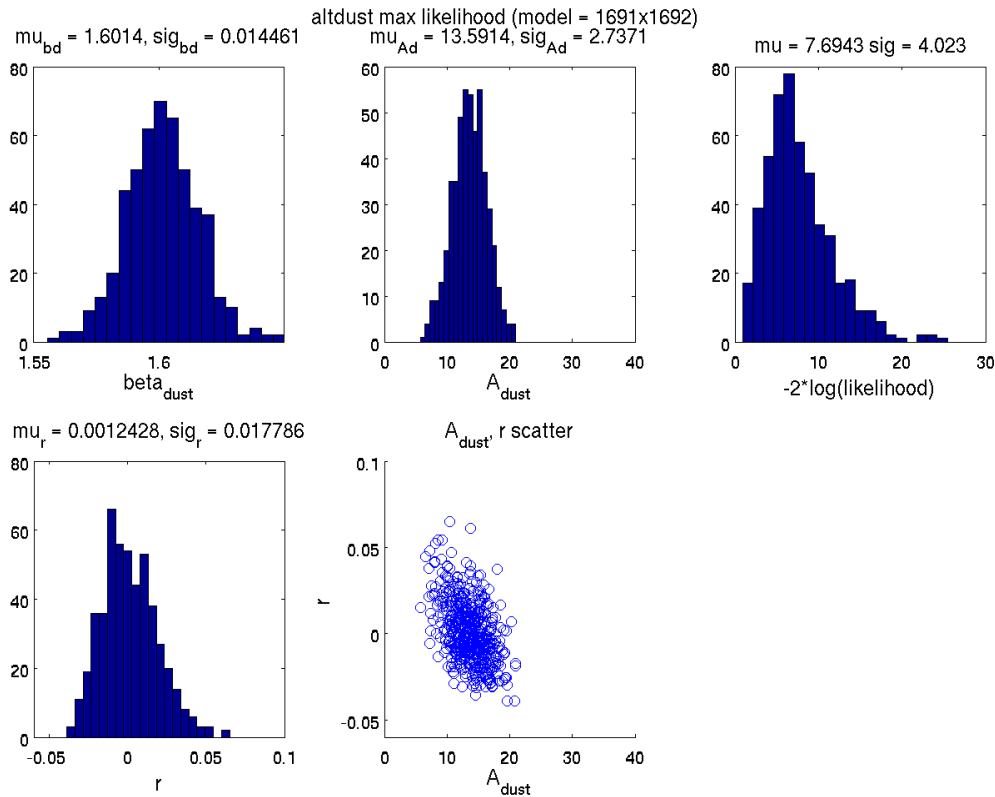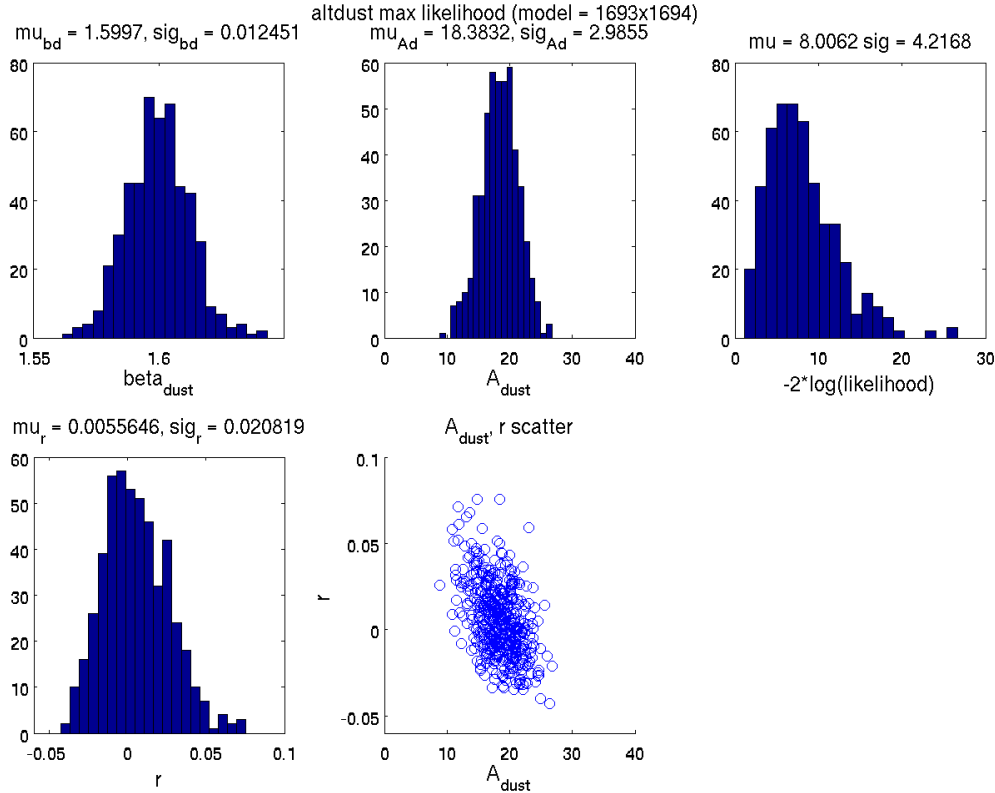
Figure 4.18: Simulations based ILC maximum likelihood search results for ILC input simulations composed from 1693×1694 altdust simulations. The top row going from left to right are: histogrammed values of β_d at point of maximum likelihood for all realizations, histogrammed values of A_d, and χ^2 of maximum likelihood values. The bottom row from left to right are the histogrammed values of r and a scatter plot of the A_d and r points for each realization at which the maximum likelihood occurs.

the nominal value 1.6, which we can attribute to the tight β_d prior.

Overall we find looser constraints when running the altdust simulations through our simulations based ILC analysis than we find with the standard simulations, an unsurprising result for several reasons. First, we place a tight constraint on β_d with a narrow prior centered at 1.6, which may not be the true value of β_d in these simulations due to the complicated nature of their foreground composition. Another complication in performing this analysis on altdust simulations is that the simulations ILC model is composed directly from noise and foreground simulations that are components of the standard simulations. The altdust simulations have a wide variation of foregound composition, and further a wide variety of

BK internal Altdust ML fg Results, no β_d prior		
Model	A_d	β_d
1682x1683	31.39	1.56
1688x1689	2.93	1.53
1691x1692	13.26	1.62
1693x1694	13.63	1.45

Table 4.3: Results from an internal BICEP/Keck 3 maximum likelihood search performed on altdust simulations [44] compared to simulations based ILC maximum likelihood search results. The ML searches performed here did not include a prior on β_d, allowing the value of this parameter to be found outside of its usual tight constraint. The findings from these results were used in the calculation of the Saha model noise and foreground term for altdust simulations.

modeling of the individual foreground components. We only have altdust simulations with all components included and do not have convenient altdust noise only or foreground only maps, and thus do not have the ability to adjust the noise and foreground biases calculated for the ILC model to accomodate the foreground composition in the altdust simulations. The ILC model is built in such a way that components other than dust are not considered, and the ML search probes only a foreground parameter space that excludes everything but dust.

While the ML search results using the simulations based ILC model are quite promising for the standard simulations from the BK15 data release, we find that this model cannot reasonably handle realizations of sky that contain different foreground composition than the standard simulations. Due to the necessity of creating simulations spreading all of parameter space, the simulations based ILC method becomes much more complicated once we think of introducing newer data, including more foregorund components, exploring a wider allowed range for A_d, or any other matter of modification. Further, this method requires the existence of simulations that contain only noise or only foregrounds in addition to simulations containing full sky data containing CMB+noise+foregrounds. Rather, it makes more sense to continue the analysis with a more general model that can quickly be adapted to the specific cases for which we take interest.

4.5   Saha 1D r Likelihood Calculation

Rather than conducting a multivariate maximum likelihood search for the Saha model, I do a one dimensional r likelihood calculation. To do this we want a model only explicitly dependent on r, so replace the noise and foreground dependent term ($S2$) in equation (4.5) with its expectation value $\langle S2 \rangle$ where we marginalize over foreground parameter space.

$\langle S2 \rangle$ **Calculation**

We use the general expectation value calculation in equation (4.49) for $\langle S2 \rangle$, first considering an analysis with foregrounds containing only dust.

$$\langle S2 \rangle = \frac{\sum_{A_d, \beta_d} S2(\beta_d, A_d) P(A_d) P(b_d)}{\sum_{A_d, \beta_d} P(A_d) P(\beta_d)} \tag{4.49}$$

$S2$ is the Saha noise and foreground term, or the second term in equation (4.5), $P(A_d)$ is the prior on A_d, and $P(\beta_d)$ is the prior on β_d given by equation (4.48).

As mentioned previously we don't have a precise measurement on the dust amplitude and need a prior to give us a reasonable range. Previously in the sims ILC model we had simply constrained the search space of the ML search algorithm, however here we will consider a second option. What we are calling the A_d prior is actually the likelihood curves of a one dimensional A_d likelihood calculation done on the Planck 353 GHz dust maps. This calculation is the H-L likelihood with simulation bandpowers coming from the standard 353 GHz simulations, and fiducial bandpowers coming from the BICEP/Keck multicomponent model. We choose the Planck 353 GHz maps as opposed to BK150 even though they are noisier because the CMB is negligible at such a high frequency. Because this frequency is the turning point for β_d, A_d and β_d are essentially decoupled and there is no need to account for any β_d in the A_d prior. We have one A_d likelihood per realization (shown in figure 4.19 so there is a unique $\langle S2 \rangle$ for each realization going into the one dimensional r likelihood
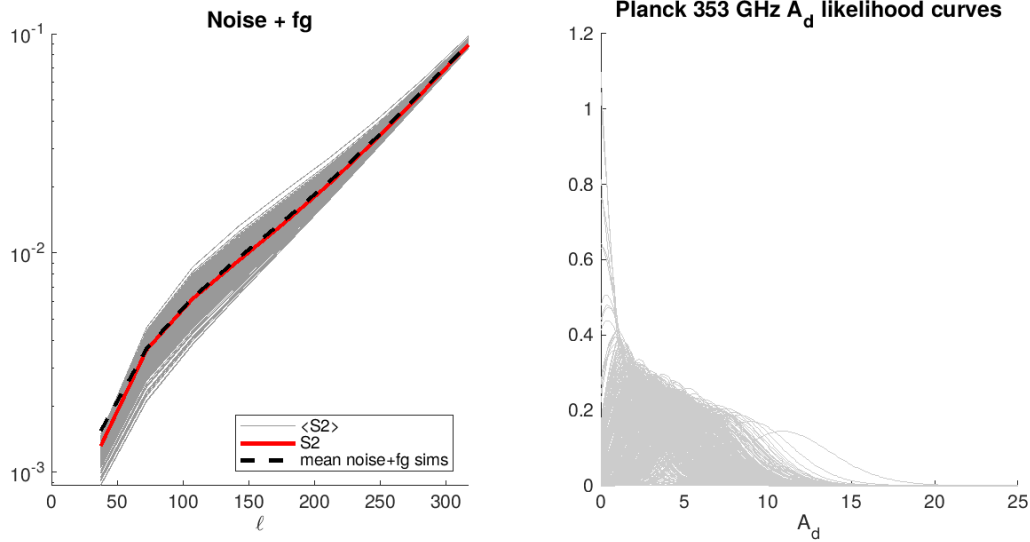
Figure 4.19: Comparing mean of noise + dust + 2*noise × dust simulations in dashed black to $S2$ in red and $\langle S2 \rangle$ curves in dark grey. Note that the red and dashed black lines are the same as in the center panel of figure 2.2, and the dark grey lines are $\langle S2 \rangle$ calculated for each realization. Planck 353 GHz auto spectrum likelihood as a function of a one dimensional model with A_d as the only free parameter is plotted to the right. Note while curves are only plotted out to $A_d = 25$, they are actually calculated out to $A_d = 50$.

calculation.

Figure 4.19 shows the $\langle S2 \rangle$ curves as the dark grey lines compared with the model term $S2$ in red, and the mean of ILC noise + dust auto spectra + 2 · noise × dust cross spectra simulations in the dotted black line. When running the H-L likelihood, each realization will have its own $\langle S2 \rangle$ curve so that the cleaned power spectrum changes from realization to realization.

Accounting for $\langle S2 \rangle$ Uncertainty

There is some uncertainty in the $\langle S2 \rangle$ term because of the noise and foreground just like in the ILC power spectrum ($\langle \hat{D}_\ell^{Clean} \rangle$) which we will account for with the variance of $S2$.

The uncertainty in using $\langle S2 \rangle$ is accounted for with the calculation of the covariance of $S2$. As mentioned in section 4.2, the H-L likelihood function requires a variable M_f, or the bandpower covariance matrix of the ILC realizations. Just as there is an $\langle S2 \rangle$ curve for each
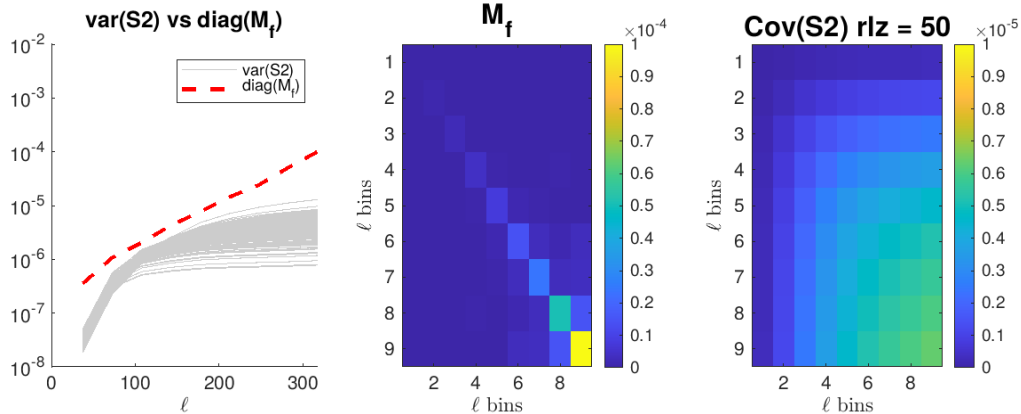
Figure 4.20: Covariance and variance of $S2$ compared with the bandpower covariance matrix of ILC simulations, M_f. In the left-hand panel are plotted the diagonals of the M_f matrix as the dashed red line, and the grey lines are the variance of $S2$ for each realization, with ℓ values on the x-axis. In the center panel is the M_f matrix. To the right is the covariance matrix of $S2$ for one realization.

realization, there is also a unique covariance matrix which is added to M_f each time the H-L likelihood is calculated. The left panel in figure 4.20 compares the diagonal of M_f in red with the variance of $S2$ in grey. The center panel shows the M_f matrix and to the right is the covariance matrix of $S2$ for one simulation. Recall that the elements more than one off of the diagonal of M_f are zeroed out because there is expected to be negligible correlation between ℓ bins that are not adjacent. The covariance of $S2$ does have nonzero bins off the diagonal, because we are making the assumption that dust follows a power law spectrum with respect to ℓ.

Monte-Carlo Markov Chain $\langle S2 \rangle$ Calculation

The method in section 4.5 requires calculating $\langle S2 \rangle$ a grid of A_d and β_d values, which can quickly become time intensive once more parameters, such as A_s and β_s are considered. It works sufficiently for standard simulations assuming only dust foreground composition, however once we intend to include synchrotron this method becomes infeasible. Instead, we introduce a Monte-Carlo method in which we sample a parameter space covering the foreground parameters given by their prior distributions.
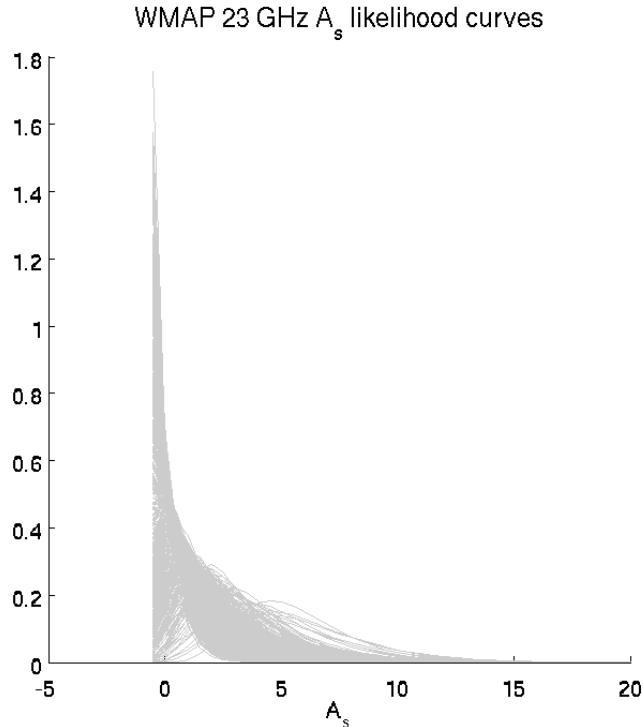
Figure 4.21: 23 GHz one dimensional A_s likelihood curves. 23 GHz is the pivot frequency for synchrotron, and the likelihood calculated at this frequency for A_s provides a reasonable region for which we can marginalize A_s.

The A_d likelihood curves and β_d prior are the same as those from section 4.5. Once synchrotron is considered, the priors for A_s and β_s are similar to those of the dust parameters. Similarly to A_d, one dimensional likelihood curves are calculated for A_s for each realization using the WMAP 23 GHz map, shown in figure 4.21. The β_s prior distribution is a Gaussian curve with $\mu_{\beta_s} = -3.1$ and $\sigma_{beta_s} = 0.3$. The parameter space from which the samples are drawn is the posterior calculated in equation (4.50). We consider the two cases of dust only, or synchrotron "turned off", and foregrounds containing both dust and synchrotron, or synchrotron "turned on". For synchrotron turned off we simply drop the A_s likelihood and β_s prior from this product.

$$posterior = P(\beta_d) \cdot P(\beta_s) \cdot \mathcal{L}(A_d) \cdot \mathcal{L}(A_s) \tag{4.50}$$
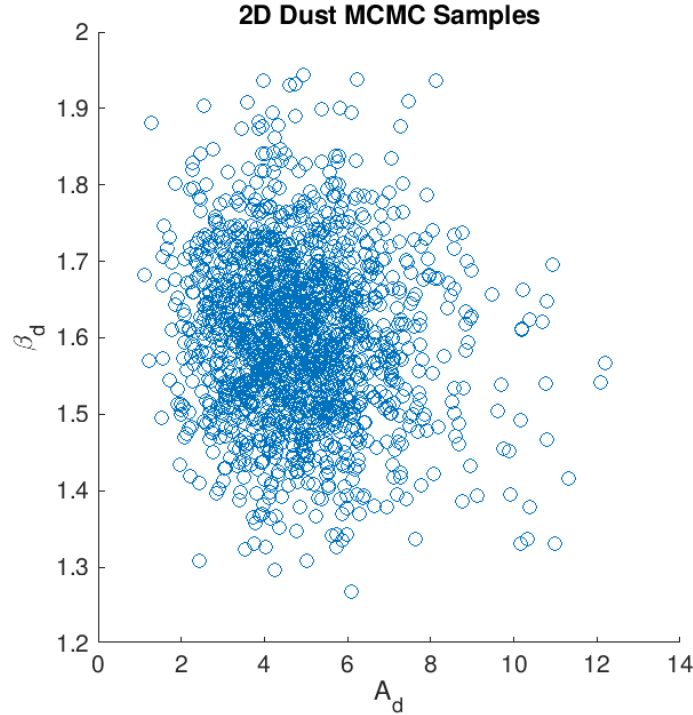
Figure 4.22: We drew 2500 samples from a region of dust parameter space described by the posterior $P(\beta_d) \cdot \mathcal{L}(A_d)$ for one of our realizations. If we were to histogram the points in the β_d direction we would recover a distribution close to that of $P(\beta_d)$, and in the A_d direction would get a distribution close to that of the corresponding simulation A_d likelihood curve.

Once we have sampled this space we calculate $S2$ at each point and take the average to get the values of $\langle S2 \rangle$. We repeat this for each realization to get a total of 499 $\langle S2 \rangle$ curves. The variance and covariances needed as described in section 4.5 are the variance and covariance of the $S2$ values calculated at each of the samples. Figure 4.22 shows a scatter plot of the samples for one realization chosen in a configuration in which the foregrounds only contain dust.

Figure 4.23 shows the $\langle S2 \rangle$ curves for each realization calculated using the MCMC method for dust only foregrounds compared with the noise and foreground simulation average and $S2$ calculated for $A_d = 3.75$ and $\beta_d = 1.6$. Figure 4.24 shows on the right a covariance matrix from one of the realizations compared to the center panel with the grid-based covariance matrix for the same realization. Plotted in the left panel are the variance curves for $S2$ with
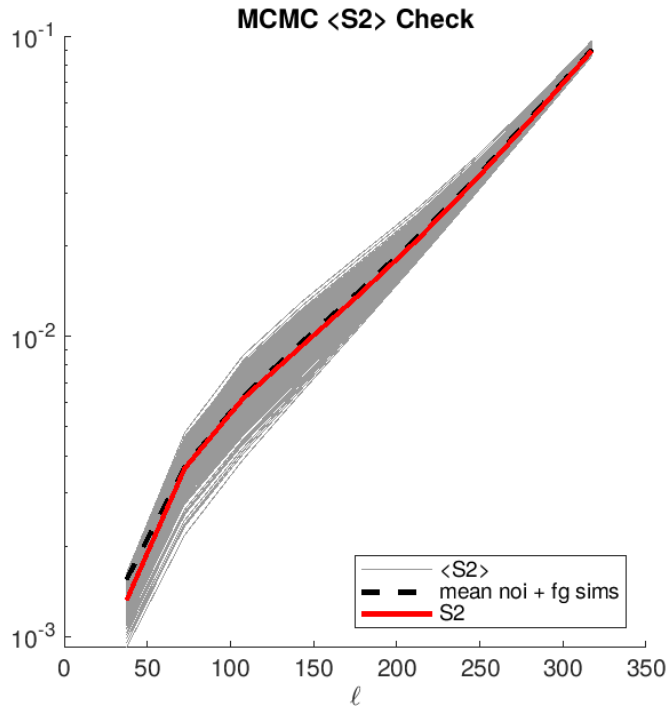
Figure 4.23: Comparing mean of noise + dust + 2*noise × dust simulations in dashed black to $S2$ in red and $\langle S2 \rangle$ curves in dark grey. Note that the red and dashed black lines are the same as in the center panel of figure 4.9, and the dark grey lines are $\langle S2 \rangle$ calculated for each realization using the MCMC sampling method. Results are the same as grid calculation shown in the left panel of figure 4.19.

the diagonal of M_f being shown as the dashed red line. We see that we obtain the same results when comparing these plots to figures 4.19 and 4.20, and continue with the likelihood calculations henceforth using the MCMC $\langle S2 \rangle$.

4.6   Saha 1D r Likelihood Results Standard Sims

The Saha likelihood performed on standard simulations with free parameter r is plotted in figure 4.25 for both synchrotron turned off in the top row and on in the bottom row. The grey curves in the left hand panels are the 499 likelihoods while the black curve is the product of the grey. The center panels are the black curves from the left panels, zoomed in on the x-axis, with a blue vertical line plotted at $r = 0$. The right hand panels are the
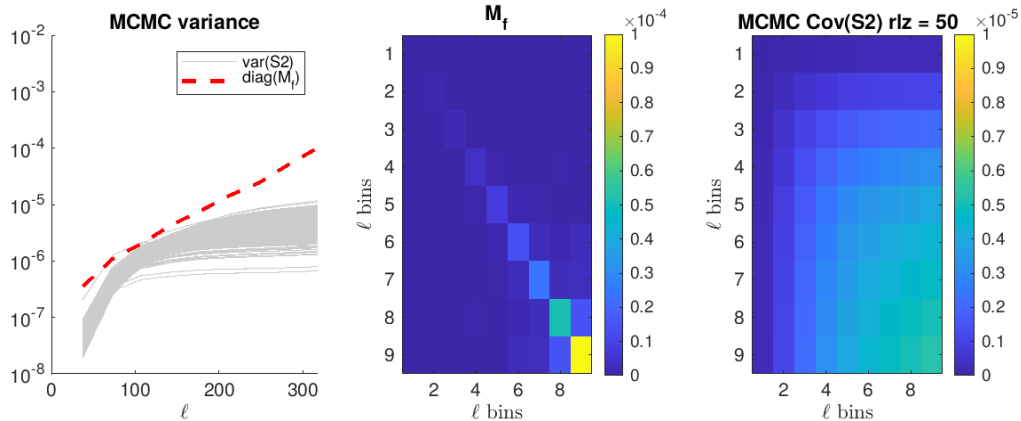
Figure 4.24: Covariance and variance of $S2$ calculated with the MCMC sampling method compared with the bandpower covariance matrix of ILC simulations, M_f. In the left-hand panel are plotted the diagonals of the M_f matrix as the dashed red line, and the grey lines are the variance of $S2$ for each realization, with ℓ values on the x-axis. In the center panel is the M_f matrix. To the right is the covariance matrix of $S2$ for the same realization plotted in figure 4.20.

histogrammed peaks of the grey curves from the left hand plots. Here we can see for both synchrotron included and left out that $\sigma_r = 0.025$, a bit larger than $\sigma_r = 0.020$ from the BK15 ML search results [34], however the bias on r is negligible, $\mu_r = -0.005$. We note that while the shapes change very slightly for synchrotron turned on versus off, the statistics of the maximum likelihood search are virtually identical for the two cases. Table 4.4 shows the ML search results for the BK15 analysis for standard sims (labeled "Gauss" in the top row) and altdust simulations. The second column shows the results in the format $\sigma_r, \mu_r/\sigma_r$, where we can interpret for the standard sims that $\mu_r/\sigma_r = 0.1\sigma$ and therefore $\mu_r = 0.002$, a result similar in magnitude to our Saha ML search result.

4.7 Saha 1D r Likelihood Results Altdust Sims

The altdust simulations were run through the ILC Saha likelihood analysis as a test of the robustness of the ILC Saha model. One change was implemented in this likelihood study compared to that of the standard simulations where the β_d prior center was adjusted for

BK15 ML Results	
Model	σ_r, $\bar{r}/\sigma_r$
Gauss	0.020, $+0.1\sigma$
1682x1683	0.028, $+0.1\sigma$
1688x1689	0.02, $+0.2\sigma$
1691x1692	0.026, $+0.2\sigma$
1693x1694	0.03, $+0.4\sigma$

Table 4.4: BK15 standard and aldust maximum likelihood search results for parameter r. The ML search performed to obtain these results used the same β_d prior used in our ILC analysis, finding tighter constraints on r with the standard Gaussian dust simulations in the top row as compared wtih the altdust simulations in the following rows.

Figure 4.25: r likelihood results for standard simulations. The grey lines are the individual simulations where the curves are normalized to integrate to 1, and the black curve is the product of the grey curves. The center panel shows the close up of the black curve from the left panel with $r = 0$ plotted in blue. The right panel shows the distribution of the r values at which the maximum likelihoods occur. The top row shows results for no synchrotron, and bottom shows results for synchrotron turned on. While there are slight differences in shape, statistics of the maximum likelihood searches are equivalent.

the different types of altdust simulations. Because the altdust foreground models are so different, we want the priors allow for a search in the appropriate parameter space. The A_d and A_s likelihoods take this into account by nature of their calculation, and we choose the center of the β_d prior from a previous internal MLsearch on altdust simulations from Ref. [44] (with results shown in table 4.3). The results for the β_s maximum likelihood search are similar for all altdust models and remains the same for all standard and altdust simulations. Figures 4.26 through 4.29 show likelihood curves and ML distributions for the different altdust models, where we have done calculations both including and excluding synchrotron component. A summary of the ILC altdust r likelihood results is compiled in table 4.5, and in the second left column are the values μ_{β_d} for each altdust model. We keep the width of the β_d prior the same for all simulations, where $\sigma_{\beta_d} = 0.11$.

We see in figure 4.26 the results for the 1682×1683 simulations, with a β_d prior centered at 1.56, a full 0.36σ shift from the nominal $\mu_{\beta_d} = 1.6$. We see in the center panels for both synchrotron excluded (top) and included (bottom) that the product of all normalized grey curves in the left panel is centered quite far from $r = 0$, where this point is not even included in the full curve. σ_r remains the same for either case, however including synchrotron moves the mean of the ML distribution closer to zero by a factor of 2. This decrease is expected, as we know that these simulations include synchrotron. The wider σ_r than the standard simulations ($\sigma_r = 0.025$) may be due to the fact that AME is included in these simulations, where our analysis does not consider this component.

Figure 4.27 shows the results for altdust model 1688×1689, which we recall has correlated dust and synchrotron of non-Gaussian shape. The β_d prior for this calculation is $\mu_{\beta_d} = 1.53$. We find a narrower $\sigma_r = 0.025$ compared to the 1682×1683 model, though again the product of the r likelihood curves is centered far from $r = 0$, which may be due to the length of the likelihood curves stretching out to higher r.

Figure 4.28 shows the results for the altdust model 1691×1692, which model β_d as a
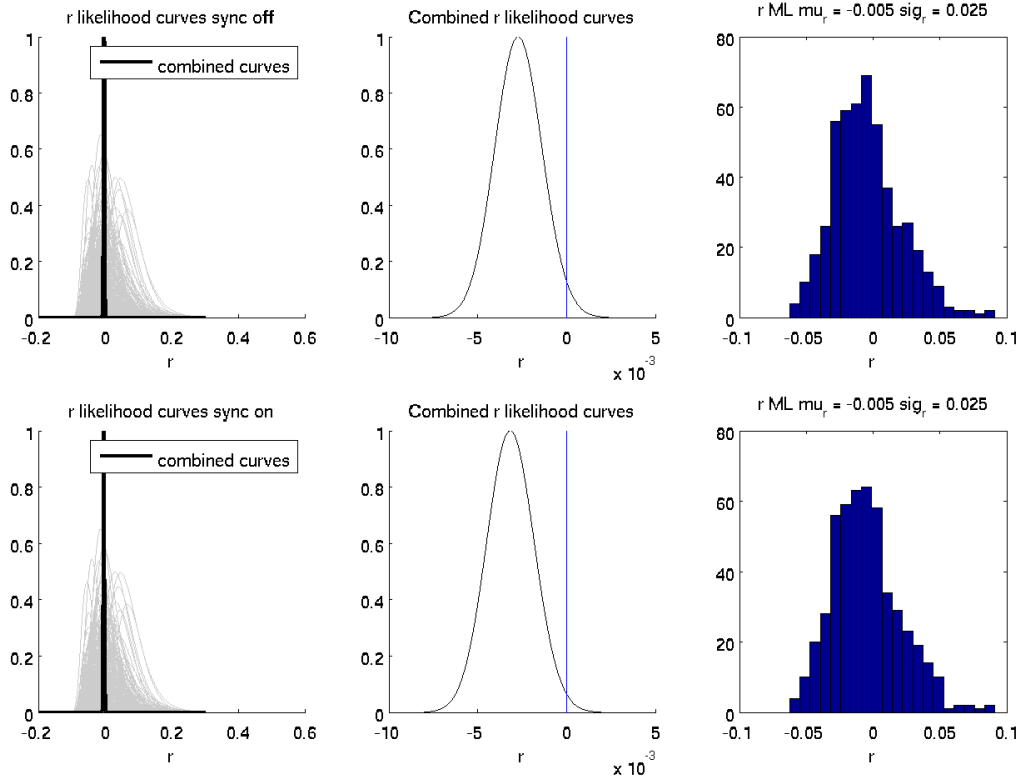
Figure 4.26: r likelihood results for 1682×1683 altdust simulations. The grey lines are the individual simulations where the curves are normalized to integrate to 1, and the black curve is the product of the grey curves. The center panel shows the close up of the black curve from the left panel with $r = 0$ plotted in red. The right panel shows the distribution of the r values at which the maximum likelihoods occur. The top row shows results for no synchrotron, and bottom shows results for synchrotron turned on.
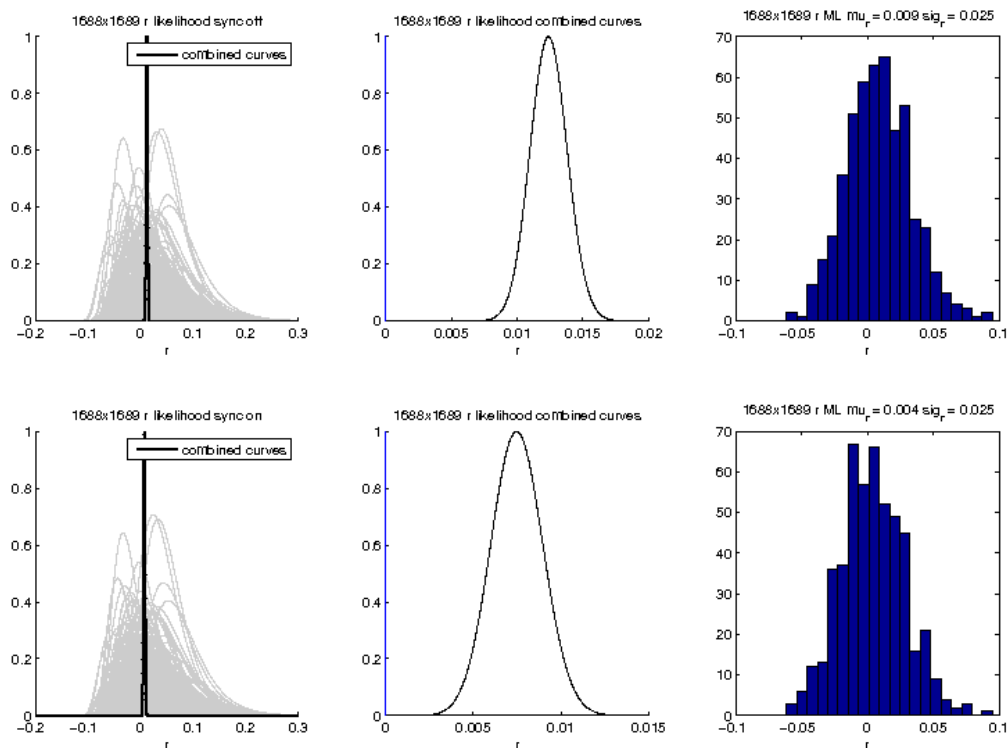
Figure 4.27: r likelihood results for 1688×1689 altdust simulations. The grey lines are the individual simulations where the curves are normalized to integrate to 1, and the black curve is the product of the grey curves. The center panel shows the close up of the black curve from the left panel with $r = 0$ plotted in blue. The right panel shows the distribution of the r values at which the maximum likelihoods occur. The top row shows results for no synchrotron, and bottom shows results for synchrotron turned on.
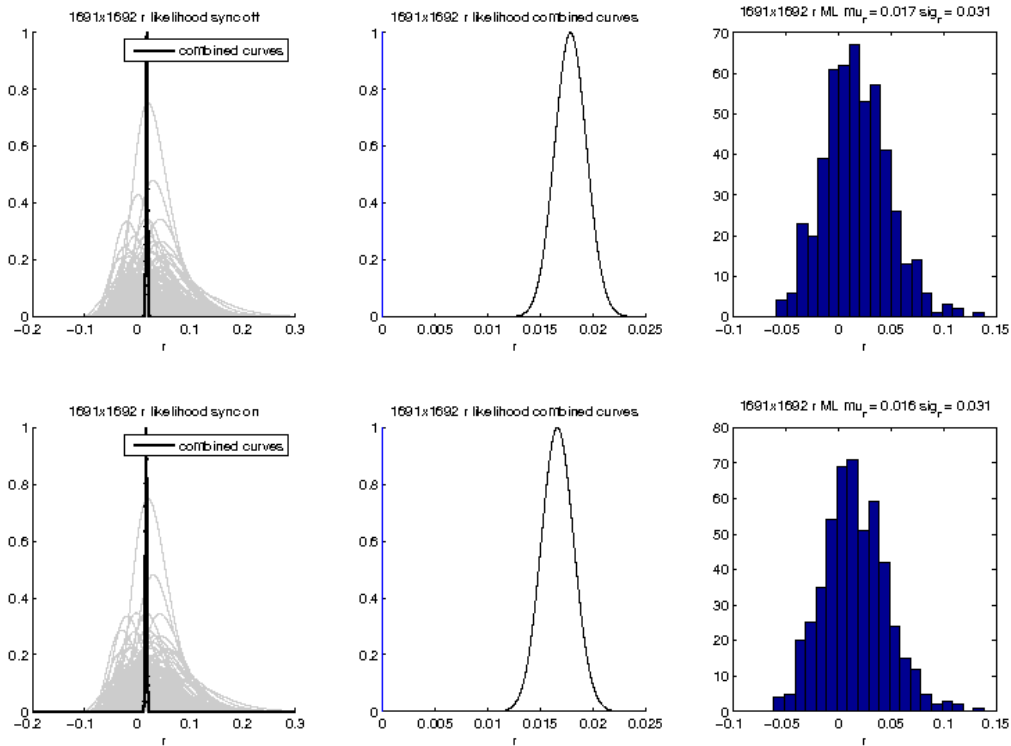
Figure 4.28: r likelihood results for 1691×1692 altdust simulations. The grey lines are the individual simulations where the curves are normalized to integrate to 1, and the black curve is the product of the grey curves. The center panel shows the close up of the black curve from the left panel with $r = 0$ plotted in blue. The right panel shows the distribution of the r values at which the maximum likelihoods occur. The top row shows results for no synchrotron, and bottom shows results for synchrotron turned on.

spatially variable parameter. The β_d prior is centered at $\mu_{\beta_d} = 1.62$, quite close to the nominal value $\beta_d = 1.6$ that is in the standard dust simulations. The results for these simulations are quite biased, where $\mu_r = 0.5\sigma$ whether synchrotron is included or not. This may have something to do with the complicated modeling of dust in the simulations, and it could be that placing a tight β_d constraint (even if it is possibly centered closer to the true value of β_d in the simulations). The product of the grey curves is quite far from $r = 0$, evident even when looking at the curve for full range of r in the left hand panels.

Figure 4.29 shows the results for the 1693×1694 simulations, where the foregrounds are a complicated SED. The likelihood curves are all normalized so that the area under each
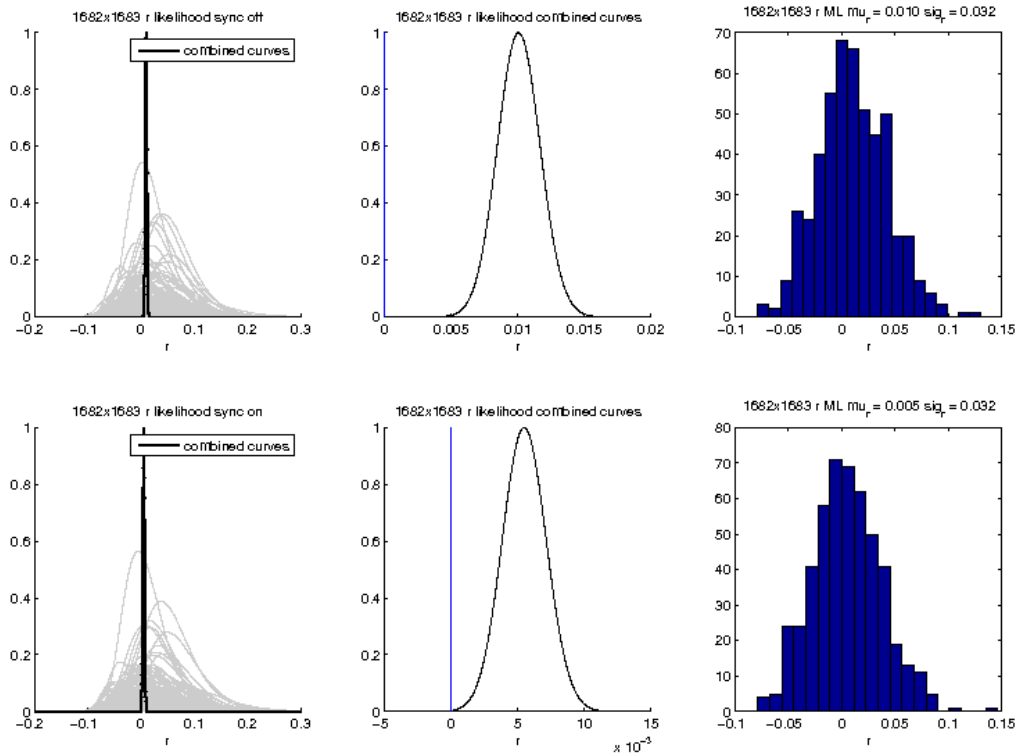
119
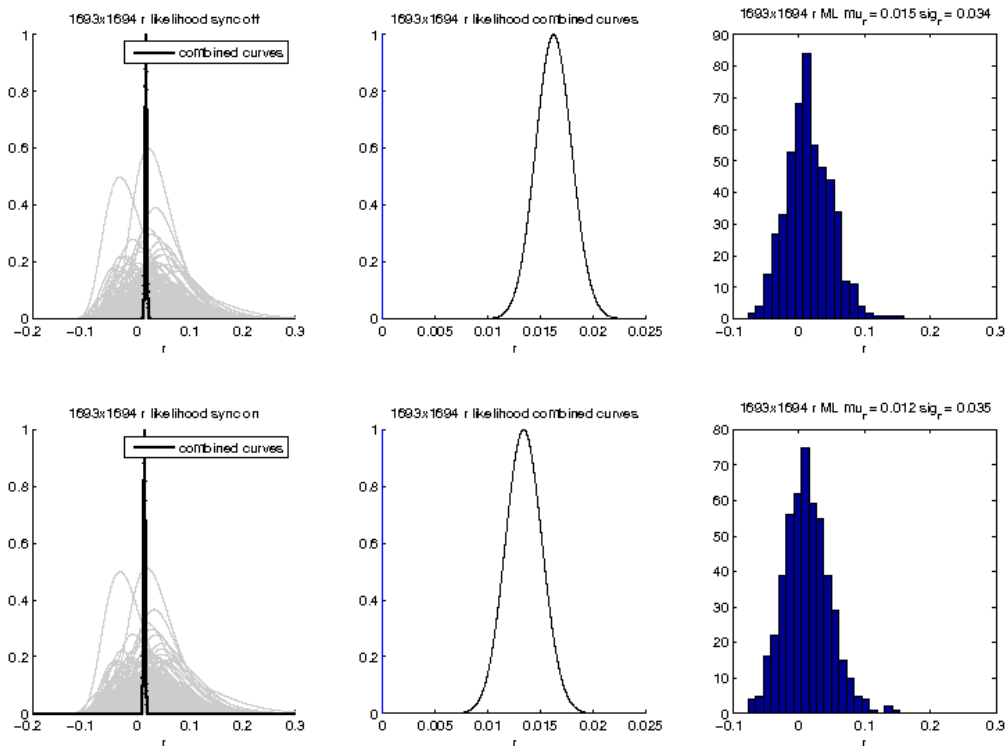
Figure 4.29: r likelihood results for 1693×1694 altdust simulations. The grey lines are the individual simulations where the curves are normalized to integrate to 1, and the black curve is the product of the grey curves. The center panel shows the close up of the black curve from the left panel with $r = 0$ plotted in blue. The right panel shows the distribution of the r values at which the maximum likelihoods occur. The top row shows results for no synchrotron, and bottom shows results for synchrotron turned on.

curve is unity. We can see the heights of the curves vary quite a bit, telling us the widths of the likelihood curves vary from realization to realization. We can see the product of the curves is nowhere near $r = 0$, and our ML results for r are quite biased with $\sigma_r = 0.034$ (or $\sigma_r = 0.035$ for synchrotron included).

The statistics of these maximum likelihood searches generally show better results (lower σ_r or less bias) should synchrotron be included in the calculation, which is unsurprising considering all altdust models have a synchrotron component. With the exception of the 1691×1692 model, we obtain a consistent level of sensitivity to r with a similar level of σ_r and bias regardless of the modeling of the foreground components. Comparing to the BK15

ILC Altdust Models ML Results			
		σ_r, $\bar{r}/\sigma_r$	
Model	β_d Prior Center	Sync off	Sync on
1682x1683	1.56	0.032, $+0.3\sigma$	0.032, $+0.2\sigma$
1688x1689	1.53	0.025, $+0.4\sigma$	0.025, $+0.2\sigma$
1691x1692	1.62	0.031, $+0.5\sigma$	0.031, $+0.5\sigma$
1693x1694	1.45	0.034, $+0.4\sigma$	0.035, $+0.3\sigma$

Table 4.5: ILC altdust models ML search results. In the second left column are listed the values at which the β_d prior is centered for each model, the second right column shows our ML search results for r assuming no synchrotron is in the models, and in the right column are results from calculations including synchrotron component in foregrounds. While σ_r is not generally affected whether synchrotron is included or not, the results for synchrotron inclusion yield slightly less biased results, an unsurprising result considering all altdust simulations have synchrotron included.

ML search results in table 4.4, we see that our analysis yields again a slightly larger σ_r but are of the same magnitude.

4.8 Saha 1D r Likelihood Varying β_d Prior Properties

A further study we tested was analyzing what happens to the ML search results on standard simulations given varying conditions on the β_d center. Thus far, we have shown results should the β_d prior be a Gaussian distribution centered at 1.6 with width of 0.1. We have obtained acceptable results with this constraint on β_d so far for our standard simulations because the were made with a β_d value of 1.6. For the real sky, however, we do not necessarily know the true value, with increasing observations and more data over the years we may find that this value is in fact not the best fit for the sky. Because of this we are interested in obtaining robust ML search results that do not rely on a tight constraint on β_d.

We tested out three separate changes to this prior (leaving all the β_s prior and A_d and A_s likelihoods the same):

1. Varying the center of the β_d prior, keeping width fixed to 0.11

2. Varying the width of the β_d prior, keeping the center fixed to 1.6.

3. Varying the center of the β_d prior, keeping width fixed to 1.

Notice items 1 and 3 are very similar, however in case 3 we expand the width of the prior by a factor of 10. These studies are done for synchrotron both included and ignored. Tables 4.7 - 4.8 show the results for these cases, which are further discussed below.

Case 1

The β_d width is fixed to 0.11, the same width used in all ML search methods done thus far. The center of the prior distribution ranged from 1.45 to 1.7. In table 4.7 are the results for the varying values of the β_d centers, with synchrotron turned off in the center column and synchrotron turned on to the right. We note that the case where the β_d center is 1.6 is the same calculation as the previous standard simulations ML search, and we use this as a comparison to for the behavior of the other β_d center values.

Focusing first on the case where we have synchrotron turned off, we notice that the value of σ_r is virtually undisturbed as we shift the center of the β_d prior to the left and to the right of it's nominal position. We notice a change in the bias μ_r/σ_r, shifting more negative as we move the β_d prior to the left, and becoming positive as we shift the prior to the right. Unexpectedly the largest this bias becomes is when $\mu_{\beta_d} = 1.45$, a value outside of what is predicted by Planck which likely skews our results. Overall, the values of this bias are incosequential given the size of σ_r. We see similar results when we turn synchrotron, and can conclude that moving the β_d prior around does not significantly affect our ML search results.

Case 2

Here we discuss the case of fixing the β_d prior center to the nominal value of 1.6, but change the width of this prior to values ranging from 0.5 to 1. Comparing first the results for synchrotron both turned off and turned on, we do not see a significant difference. The bias

MLSearch Varying β_d Prior Center, Width = 0.11		
	σ_r, $\bar{r}/\sigma_r$	
β_d Center	Sync Off	Sync On
1.45	0.026, -0.5σ	0.027, -0.6σ
1.5	0.026, -0.3σ	0.026, -0.5σ
1.55	0.026, -0.2σ	0.026, -0.3σ
1.6	0.025, -0.2σ	0.025, -0.2σ
1.65	0.025, +0.1σ	0.025, -0.0σ
1.7	0.025, +0.2σ	0.025, +0.1σ

Table 4.6: ILC MLSearch results for varying β_d prior centers and width = 0.11 for standard simulations. In the left column are listed the range of β_d prior centers, the center column lists results for r ML search with no synchrotron included in calculation, and the left column shows results for synchrotron included. We see for either case the value of σ_r is not greatly affected by the position of hte β_d prior, however the bias on the results tends to be smaller close to the true value of $\beta_d = 1.6$ in the simulations. The bias on the results when β_d is close to this value does not change much whether synchrotron is included in the calculation or not.

is slightly higher for synchrotron turned on which is unsurprising given the fact that the standard simulations are created assuming no synchrotron, however these values are not unacceptable. The value of σ_r varies quite slightly given the wide range on the width of the β_d prior. We can conclude here that a large width on the β_d prior will not significantly affect our results. We recall the discussion in section 4.4 that due to the narrow constraint on β_d that it is not a true free parameter. The findings from this case allow us to loosen the constraint on β_d which means that we may make less assumptions about it.

Case 3

In this final case we test the ML search behavior given a width prior with $\sigma_{\beta_d} = 1$, varying the position of its center from 1.45 to 1.7, the same range done in the first case. We would expect that given our conclusions from cases 1 and 2 that this case should not significantly affect our ML search results. Looking to table 4.8 we see that this is indeed the case. For foregrounds with synchrotron turned off we see virtually no change in ML search results given the position of the β_d prior. We do slightly lose some sensitivity as σ_r goes up to 0.027 compared to our first Saha ML search result of $\sigma_r = 0.025$, however our results remain

MLSearch Varying β_d Prior Width, Center = 1.6		
	σ_r, $\bar{r}/\sigma_r$	
β_d Width	Sync Off	Sync On
0.05	0.025, -0.1σ	0.025, -0.2σ
0.11	0.025, -0.2σ	0.025, -0.2σ
0.15	0.025, -0.0σ	0.026, -0.2σ
0.2	0.026, -0.1σ	0.026, -0.2σ
0.5	0.027, -0.1σ	0.027, -0.2σ
1	0.027, -0.1σ	0.027, -0.2σ

Table 4.7: ILC MLSearch results for β_d prior varying width and center = 1.6 for standard simulations. In the left column are listed the range of σ_{beta_d}, the center column lists results for r ML search with no synchrotron included in calculation, and the left column shows results for synchrotron included. We see that widening the β_d prior by a factor of 20 only causes us to lose 8% sensitivity on r, with unbiased results for all cases. We see no change in the values of σ_r whether synchrotron is included or not, however have more biased results when synchrotron is inlcuded.

unbiased. Again unsurprisingly our results are slightly biased when synchrotron is turned on, because we allow the ML search to be done in a parameter space that we know that the simulations do not occupy.

Saha results for altdust sims varying β_d prior

In table 4.9 we see ML search results on altdust simulations where synchrotron is present and the β_d prior has width $\sigma_{\beta_d} = 1$. In the center column are the results when we allow the center of the prior to vary in the same way was was done in table 4.5, and in the right column are ML search results when our β_d prior is always centered at 1.6. We see immediately that our σ_r values do not change whether we allow the β_d prior to move or not, and comparing these results to the right column in talbe 4.5 also tells us that we do not have to pay a price when loosening the constraints on β_d.

MLSearch Varying β_d Prior Center, Width $= 1$		
	σ_r, $\bar{r}/\sigma_r$	
β_d Center	Sync Off	Sync On
1.45	0.027, -0.1σ	0.027, -0.3σ
1.5	0.027, -0.1σ	0.027, -0.3σ
1.55	0.027, -0.1σ	0.027, -0.2σ
1.6	0.027, -0.1σ	0.027, -0.2σ
1.65	0.027, -0.1σ	0.027, -0.2σ
1.7	0.027, -0.0σ	0.028, -0.2σ

Table 4.8: MLSearch results for β_d prior varying centers and $\sigma_{bd} = 1$ for standard simulations. In the left column are listed the range of β_d prior centers, the center column lists results for r ML search with no synchrotron included in calculation, and the left column shows results for synchrotron included. We see that with a wide β_d prior the center position does not affect the results. Our synchrotron excluded results are unbiased versus the biased results when we included synchrotron in calculations.

MLSearch Wide β_d Prior $\sigma_{\beta_d} = 1$		
	σ_r, $\bar{r}/\sigma_r$	
β_d Altdust Model	Moving μ_{β_d}	$\mu_{\beta_d} = 1.6$
1682x1683	0.030, -0.4σ	0.030, -0.5σ
1688x1689	0.025, +0.2σ	0.025, +0.2σ
1691x1692	0.031, -0.1σ	0.031, -0.1σ
1693x1693	0.034, +0.2σ	0.034, +0.2σ

Table 4.9: ILC altdust MLSearch results for wide β_d prior. In the right column are results when we keep $\mu_{\beta_d} = 1.6$ for all models, and the center column shows results where each model has its own μ_{β_d}, the same centers as in the second left column of table 4.5. We see that fixing μ_{β_d} for all models does not change our results, a welcome outcome because when analyzing real data we wish to make the least amount of assumptions possible.

4.8.1 Saha results for 5D posterior

Thus far we have studied several variations on the r likelihood ML search in the way that we calculate $\langle S2 \rangle$. We have varied the shape and position of the β_d prior that is used in the posterior from which we sample to calculate $\langle S2 \rangle$, and we have calculated r likelihood curves both with synchrotron inclusion and without synchrotron. We have found that although our standard Gaussian dust simulations were created without synchrotron present, our ML search results including synchrotron are the same as when we assume foregrounds only containing dust. We have also found that including synchrotron improves our ML search results for the altdust simulations.

As a reminder, the way in which we have included synchrotron in our calculations is through the posterior from which we sample for the $\langle S2 \rangle$, calculated in equation (4.50). While this yields acceptable ML search results, it is not exactly correct, as the way that we multiply the one dimensional A_d and A_s likelihoods together treats the dust and synchrotron as independent signals.

In reality, there is some correlation between these two foreground components evidenced by data from Planck due to the fact that both elements are experiencing the same Milky Way magnetic field. We are able to include this correlation by introducing a new parameter, ϵ, which describes the dust and synchrotron correlation. This parameter can take any value between -1 and 1, and we can constrain the foregrounds with a five dimensional likelihood calculation with the free parameters A_d, β_d, A_s, β_s, and ϵ, and multiply by our same β_d and β_s priors to obtain a posterior from which to sample for $\langle S2 \rangle$.

We calculate this five dimensional likelihood by looking at the cross-correlation between maps of high and low frequencies. We continue to use the WMAP 23 GHz and Planck 353 GHz maps, and again use the H-L likelihood formula. Our new 5 dimensional posterior is then
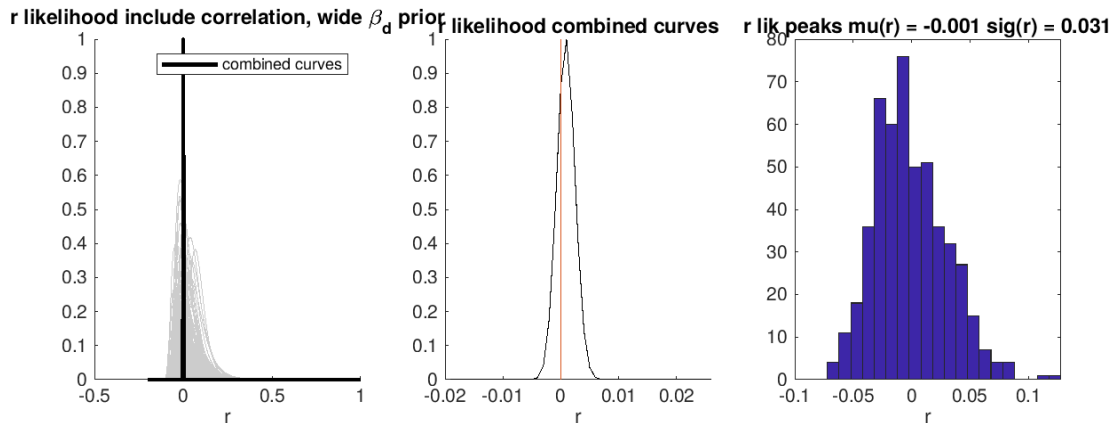
Figure 4.30: On the right are the r likelihood curves for $\langle S2 \rangle$ calculated with the 5D posterior described in equation 4.51. The grey lines are the individual simulations where the curves are normalized to integrate to 1, and the black curve is the product of the grey curves. The center panel shows the close up of the black curve from the left panel with $r = 0$ plotted in red. The right panel shows the distribution of the r values at which the maximum likelihoods occur.

$$P_{5D} = P(\beta_d) \cdot P(\beta_s) \cdot \mathcal{L}(A_d, \beta_d, A_s, \beta_s, \epsilon) \tag{4.51}$$

where $P(\beta_d)$ and $P(\beta_s)$ are the same β_d and β_s priors we have been utilizing, and $\mathcal{L}$ is our five dimensional likelihood constraining synchrotron, dust, and their correlation. We rerun our r ML search on standard and altdust simulations using this posterior with the wide β_d prior where $\sigma_{\beta_d} = 1$. The results for the standard simulations run through this configuration are shown in figure 4.30, and the altdust simulation results are listed in table 4.10.

We can see in the standard simulation maximum likelihood search that the results remain unbiased, with $\mu_r = -0.032\sigma$. When performing the maximum likelihood search on altdust simulations we find comparable σ_r, however the results (barring that of the 1691×1692 model) are quite biased with μ_r on the order of 30% to 40% of σ_r.

Altdust MLSearch 5D Posterior Sampling Distribution	
β_d Altdust Model	σ_r, $\bar{r}/\sigma_r$
1682x1683	0.031, -0.4σ
1688x1689	0.026, +0.4σ
1691x1692	0.031, -0.0σ
1693x1693	0.035, +0.3σ

Table 4.10: ILC altdust MLSearch results for 5D sampling posterior distribution. Our results do not change considerably when compared with those in table 4.9, telling us we do not loose significant sensitivity when loosening the constraints on our foreground parameter space.

Chapter 5

ILC Likelihood for Real Data

In the previous chapter we tested the effects of varying the manner in which $\langle S2 \rangle$ was calculated, ending with the method of marginalizing the sampling foreground parameter space with a 5D posterior distribution. In this chapter I will show our ML search results using this variation of the Saha model for the two cases of a narrow β_d prior where $\sigma_{\beta_d} = 0.11$, and the wide β_d prior where $\sigma_{\beta_d} = 1$. I will directly compare these results to those from the BICEP/Keck 2015 analysis. The altdust simulations previously discussed will be left out of this discussion because while they proved a useful metric of the robustness of our ILC model, the standard simulations are a better representative of the real data.

5.1   Real Data ILC Power Spectrum

We first calculate an ILC power spectrum using the real data frequency maps, and plot this in figure 5.1 as a black line with the familiar grey envelope of simulations, and the mean of sims plotted in blue. We see that the real data is well within the grey envelope, telling us it is consistent with our simulations. Once we have this power spectrum, we can then use it to calculate a likelihood curve.
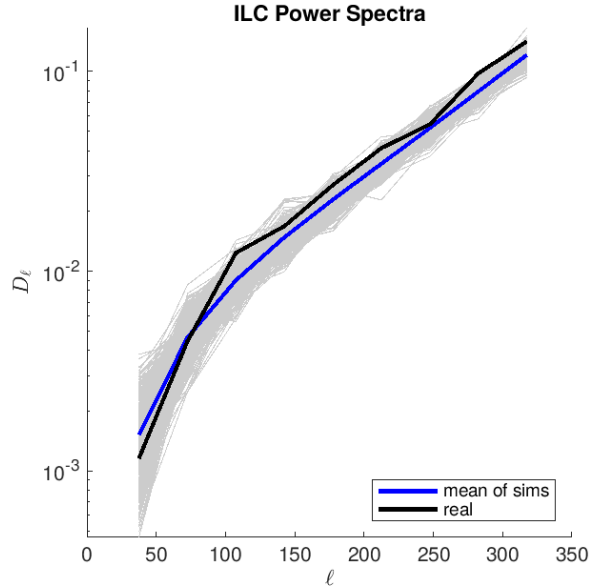
Figure 5.1: Full sky standard ILC power spectra. The grey lines are the familiar ILC simulations created with the standard frequency simulations containing Gaussian dust, the blue line is the average of the simulations, and the black line is the result of putting the real data frequency maps through the ILC process.

5.2 Real Data Likelihood

We recall that the Saha ILC model calculation depends on the term $\langle S2 \rangle$, which is calculated with an ensemble of samples drawn from a posterior distribution composed from a 5D likelihood providing a reasonable foreground parameter space. This likelihood is calculated for each realization with the WMAP 23 GHz and Planck 353 GHz bandpowers and their cross spectra, such that each realization that goes through the HL likelihood calculation has it's own $\langle S2 \rangle$ curve. For our real data likelihood calculation, we keep the same process, however calculate this 5D foreground likelihood using the real foreground data from the BK15 data release.

We multiply this likelihood by the same β_d and β_s priors from equation (4.51), and choose to analyze separately the results with a wide β_d prior versus a narrow β_d prior. Our reasoning for including a wide prior is that it allows us to consider a wider area of β_d space were we to think that the historical value of 1.6 were incorrect. We recall for example that in section

4.4, our altdust results were quite biased which we postulated was partially due to the tight β_d prior influencing results giving us a value much different than the true value in the input frequency simulations.

While all of our simulation results in this chapter will be calculated with β_d priors centered about $\beta_d = 1.6$, recent Planck results [45] have found that the best fit value for β_d may actually be closer to $\beta_d = 1.53$. Our standard simulations were created with a true value of $\beta_d = 1.6$, so we will continue to analyze them with a prior centered at this value. However we will also see what happens to the real data likelihood curve with a model where $\langle S2 \rangle$ is calculated with a narrow β_d prior centered about $\beta_d = 1.53$.

We see in figures 5.2 and 5.3 the ML search results shown in previous sections, with the addition of a curve calculated from real data. The plots in the left panels show the grey envelope of simulation r likelihoods, and a black curve showing the likelihood curve of the real data. We see that the simulation curves calculated with a wide prior have longer tales than the simulation curves calculated with a narrow prior, however the shape of the real data curve does not seem to change significantly with the prior width.

The right hand panels in figures 5.2 and 5.3 show the distribution of the r values at which the likelihoods reach their maximum. In both plots the red line comes from the BK15 analysis giving a maximum likelihood for the real data at $r_{ML,BK15} = 0.02$. The black lines are the maximum likelihood position for our ILC analysis, where for a narrow β_d prior we get a value of $r_{ML,narrow} = -0.005$ and for a wide prior is $r_{ML,wide} = -0.018$.

Figure 5.4 shows the real data likelihood curves from figures 5.2 and 5.3, with the addition of a likelihood curve calculated with the β_d prior centered about $\beta_d = 1.53$ with width $\sigma_{\beta_d} = 0.11$. When comparing the values of r for the maxima of each of the curves, we find that for the wide β_d prior centered at $\beta_d = 1.6$ plotted in black, $r_{ML} = -0.022$, which is quite close to the value when using the narrow prior centered at $\beta_d = 1.53$ plotted in blue, where we get $r_{ML} = -0.018$. The curve calculated with a narrow β_d prior centered at
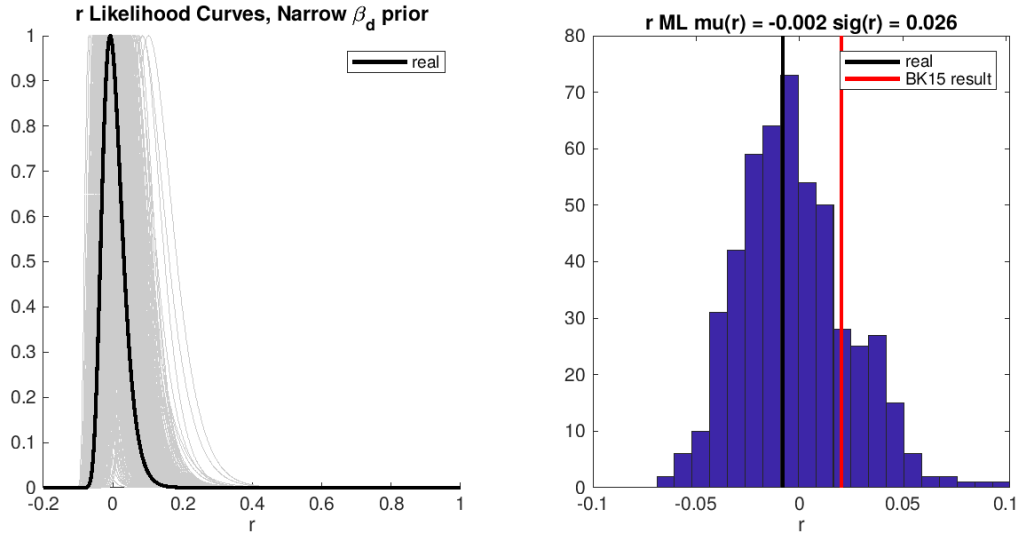
Figure 5.2: Maximum likelihood search results from our ILC analysis using standard simulations, where the $\langle S2 \rangle$ term in the Saha model used in the HL likelihood is calculated with a β_d prior with width $\sigma_{\beta_d} = 0.11$. Left: r likelihood curves where the grey lines are calculated with the 499 simulations, and the black curve is calculated with real data. Right: the r values at which the likelihood curves peak. The black line is the value from the real data put through the ILC analysis, and the red line is the BK15 result.
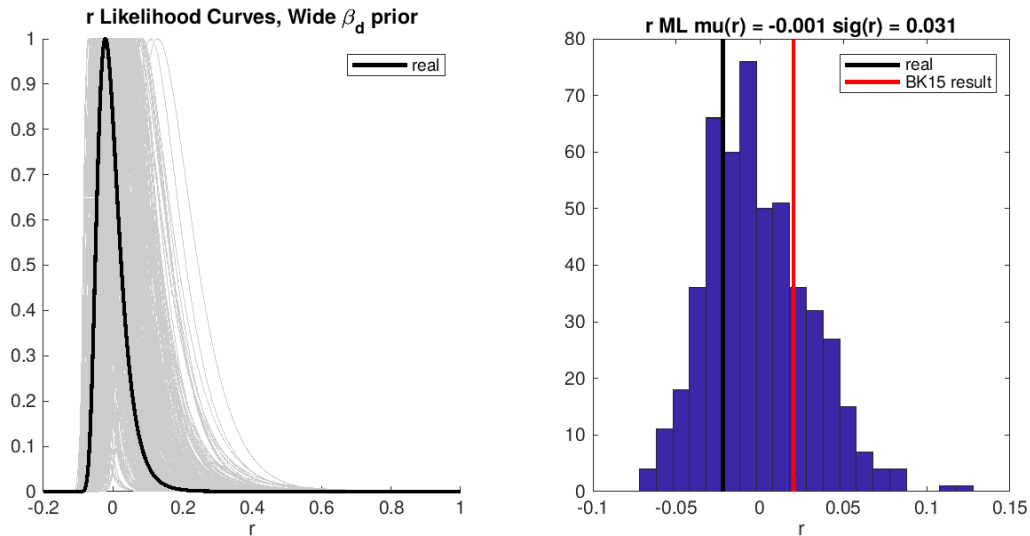

Figure 5.3: Maximum likelihood search results from our ILC analysis using standard simulations, where the $\langle S2 \rangle$ term in the Saha model used in the HL likelihood is calculated with a β_d prior with width $\sigma_{\beta_d} = 1$. Left: r likelihood curves where the grey lines are calculated with the 499 simulations, and the black curve is calculated with real data. Right: the r values at which the likelihood curves peak. The black line is the value from the real data put through the ILC analysis, and the red line is the BK15 result.
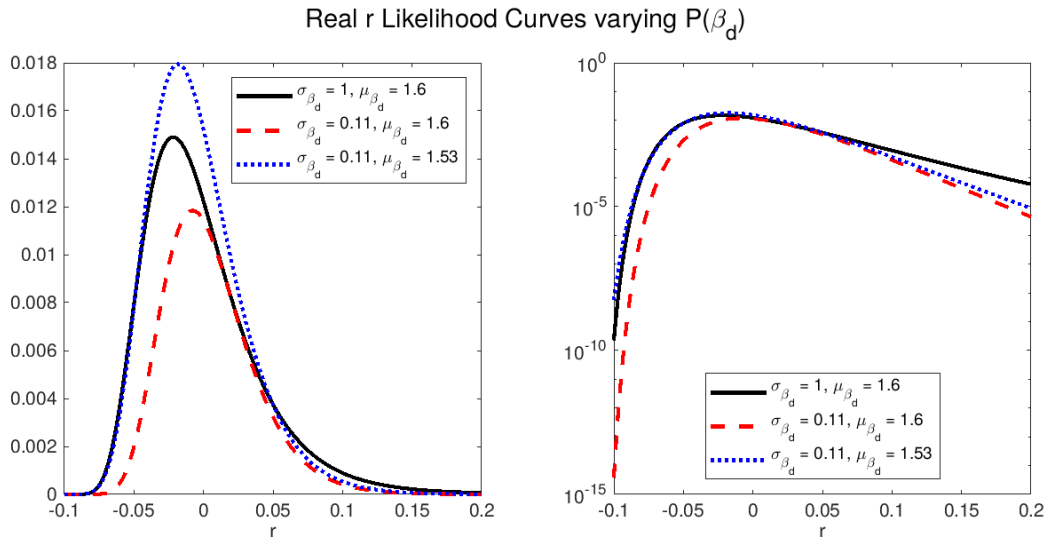
Figure 5.4: Real data likelihood curves. The black line shows the r likelihood curve where the $\langle S2 \rangle$ term was calculated with a wide β_d prior centered at 1.6, the dashed red shows the curve where $\langle S2 \rangle$ was calculated with a narrow β_d prior centered at 1.6, and the blue dotted line shows the curve where $\langle S2 \rangle$ was calculated with a narrow β_d prior centered at 1.5.

$\beta_d = 1.6$ plotted in red has a different shape when compared to the other two curves, and has $r_{ML} = -0.008$. The plot on the left shows these curves on a linear y-axis, and on the right are the same curves plotted on a log y-axis. The fact that the red curve is the most different of the three plotted shows that how we marginalize over the foreground parameters influences the value of r that our analysis finds. More specifically, the allowed region for the β_d parameter has influence on where we find our best fit for r.

Figure 5.5 shows scatter plots for the ML results for simulations put throught the ILC analysis (calculated with a narrow β_d prior in the left panel and a wide β_d prior in the right) compared to the BK15 results. The black dot on each plot shows the value of r_{ML} obtained from the real data. We see a clear correlation between the ILC analysis and the BK15 analysis results, even if they do not give quite the same result. We see stronger correlation when using a narrow β_d prior than a wide one, however neither method gives large outliers. On both panels is plotted the real r_{ML} result when calculated with a narrow β_d prior centered at $\beta_d = 1.53$, and we can see confirmation to figure 5.4 that the maximum r likelihood value
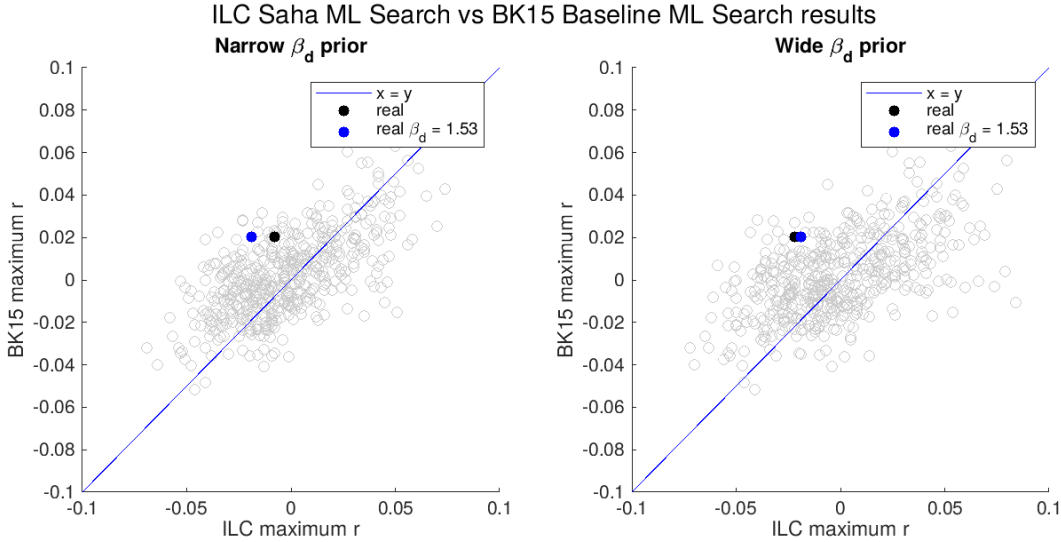
Figure 5.5: Scatter plots showing the maximum likelihood r values from our ILC analysis compared those found from the BK15 analysis. The left plot has $\langle S2 \rangle$ calculated with a narrow β_d prior, and the right panel is calculated with a wide β_d prior. The grey points in these plots are results from simulations, the black dot is the real data result, and the blue dot is the real result where the β_d prior is centered at $\beta_d = 1.53$.

is closer to the calculation with a wide β_d prior than a narrow centered at $\beta_d = 1.6$.

Another way of directly comparing our ILC analysis results realization by realization to the BK15 analysis results is plotted in Fig. 5.6. This one is a bit more complicated to interpret than the results from Fig. 5.5. The x-axis on these plots shows the difference in r_{ML} values obtained for each realization between each analysis, where we subtract the ILC analysis $r_{ML,ILC}$ value from the BK15 $r_{ML,BK15}$. On the y-axis is plotted the best fit values for β_d from the BK15 analysis. Again the grey points show simulation results, and the solid black dot is the real data. We can see a bit of correlation which is more prominent in the calculations with a narrow β_d prior. This tells us that the realizations where the BK15 ML search finds a low β_d are realizations where the ILC analysis finds a lower r value than the BK15 analysis. This could mean that the BK15 ML search tends to finds a low β_d when it finds high r. The data in each case falls within the spread of simulations, however is in the higher end of $r_{ML,BK15} - r_{ML,ILC}$. We have again in blue plotted the real data r_{ML} result for narrow β_d centered at $\beta_d = 1.53$ on both plots.
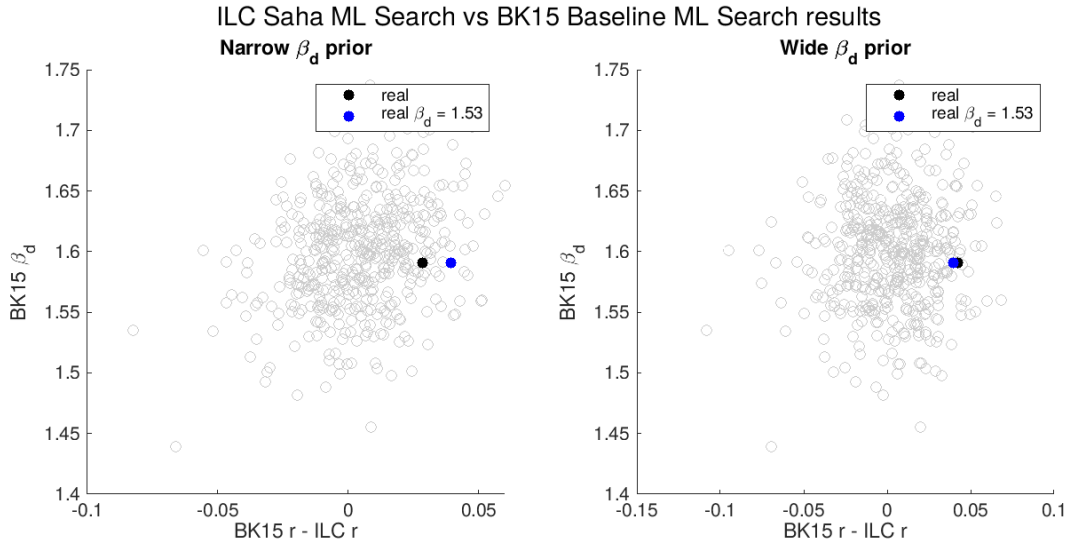
Figure 5.6: Scatter plots showing relationship between difference in BK15 and ILC analysis r ML results, and the BK15 β_d results. On the left are the points comparing BK15 to the narrow β_d prior ILC results, and on the right are the points comparing BK15 to the wide β_d prior ILC results. On both plots the blue dot is the real result where the β_d prior is centered at $\beta_d = 1.53$

5.3 Zero-to-Peak Ratio

We can calculate the ratio between the likelihood at $r = 0$ and $r = r_{ML}$, or the zero-to-peak ratio L_0/L_{peak}, as another way to analyse the distribution the r_{ML}. We only look at the posterior curves for $r > 0$, and because we know that a significant portion of the r posteriors peak at negative r, since we are cutting off that portion of parameter space we expect a significant portion of our L_0/L_{peak} values to be equal to 1. We calculate L_0/L_{peak} for all realizations, and plot the CDF of these ratios in red in the right panels of figures 5.7 and 5.8. The black line is the simple analytic ansatz $\frac{1}{2}(1 - f(-1logL_0/L_{peak}))$ where f is the χ^2 CDF for one degree of freedom. This line is what we would expect should the widths of the individual likelihood curves were a good representation of the scatter of their peaks. The left panels have replotted the likelihood curves normalized by their maxima for positive r, with the real data plotted in black. Figure 5.9 shows these same plots from the BK15 paper.

We see that the BK15 analysis found a positive maximum value for r_{ML} in the real data,

and have plotted in the dashed black line where this occurs on their L_0/L_{peak} plot. They find that 19% of the simulations run through their analysis have a zero-to-peak ratio less than the real data, meaning that these realizations show more evidence for a positive value for r when the true value is in fact zero (we recall that the simulations are created with $r = 0$). 53% of their simulations have zero-to-peak ratio of 1, meaning that these simulations have a maximum likelihood value occurring at $r \leq 0$.

Our ILC analysis shows that 40% (44% for a wide β_d prior) of our simulations have zero-to-peak ratio less than the real data, which we recall is equal to 1 because we find that the real data likelihood curve peaks at negative r. Ideally our CDF would agree with the analytic ansatz, which the results of BK15 does quite well, however we find quite the gap between our CDF and this ansatz, indicating that the peaks of our likelihoods are more tightly clustered towards $r = 0$ than we might expect by looking at the individual likelihood curves. In other words, the likelihood curves are a bit wider than we would expect by looking at the cluster of peaks.

5.4   Highest Density Posterior Interval

While it is very important to consider data maximum likelihood results, we can also find further information when looking at the highest posterior density interval, which is the narrowest interval for some given confidence interval. We have calculated this for the 68% confidence interval, and the regions calculated for narrow (with priors centered at $\beta_d = 1.6$ and $\beta_d = 1.53$) and wide β_d priors is shown in figure 5.10 in the left, right, and center panels respectively. The posteriors plotted are the real data likelihood curves normalized such that the area under each curve is 1. The orange shaded regions show the 68% HPD interval, and we can see that the width of the regions for the various calculations is fairly similar, though shifted more negative for the center and right panels as compared with the left panel. The HPD calculated from the narrow prior curve centered at $\beta_d = 1.53$ is $r = -0.008^{+0.034}_{-0.025}$, from
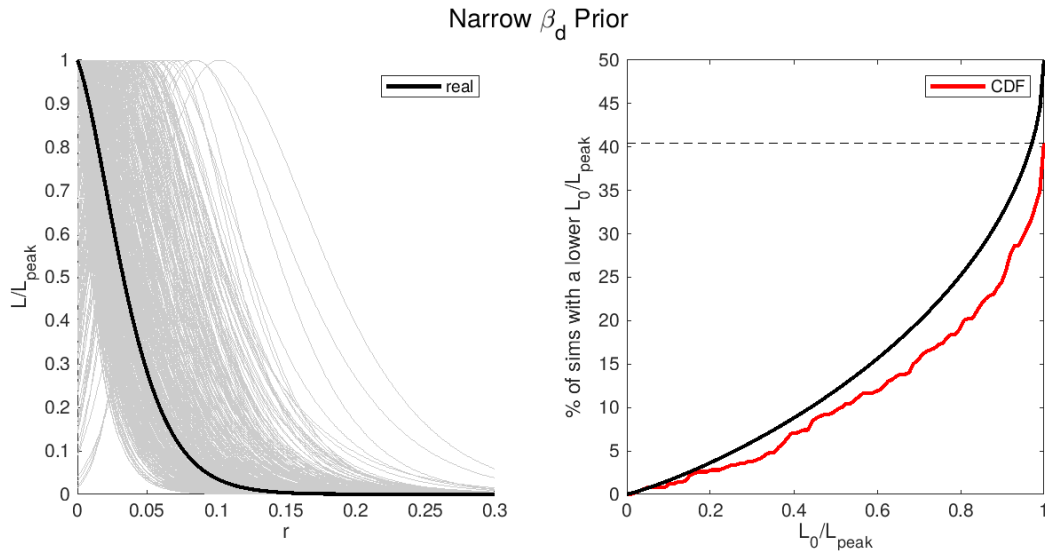
Figure 5.7: Zero-to-peak results when calculated with a narrow β_d prior. Left: r likelihoods normalized by their maximum likelihood values for the region $r \geq 0$, where grey lines show simulations and the black line is the real data result. Right: CDF of the zero-to-peak ratio (value of the likelihood at $r = 0$ divided by it's maximum value) in red, and the simple analytic ansatz $\frac{1}{2}(1 - f(-2logL_0/L_{peak}))$ plotted in black, where f is the χ^2 CDF for one degree of freedom.

Figure 5.8: Zero-to-peak results when calculated with a wide β_d prior. Left: r likelihoods normalized by their maximum likelihood values for the region $r \geq 0$, where grey lines show simulations and the black line is the real data result. Right: CDF of the zero-to-peak ratio (value of the likelihood at $r = 0$ divided by it's maximum value) in red, and the simple analytic ansatz $\frac{1}{2}(1 - f(-2logL_0/L_{peak}))$ plotted in black, where f is the χ^2 CDF for one degree of freedom.
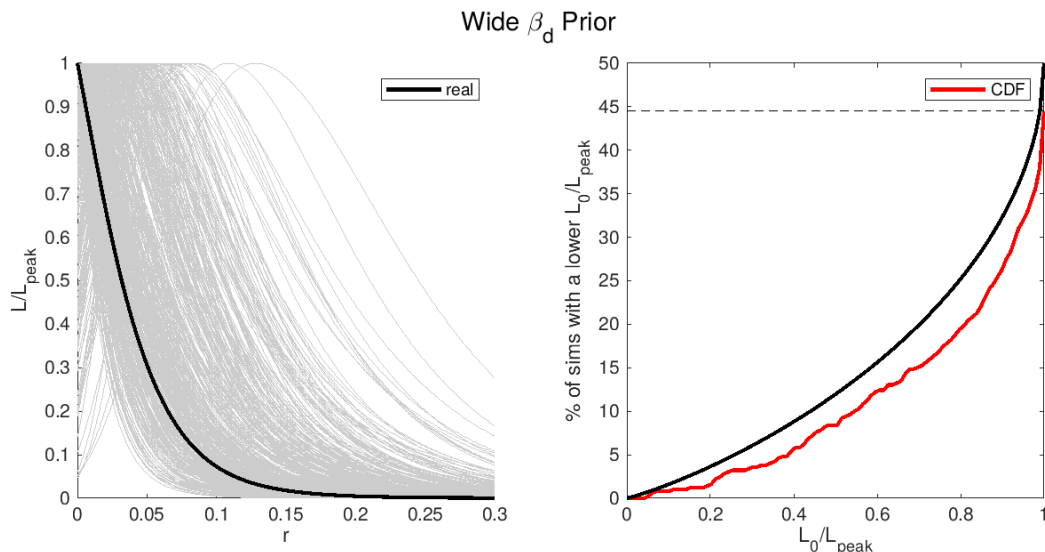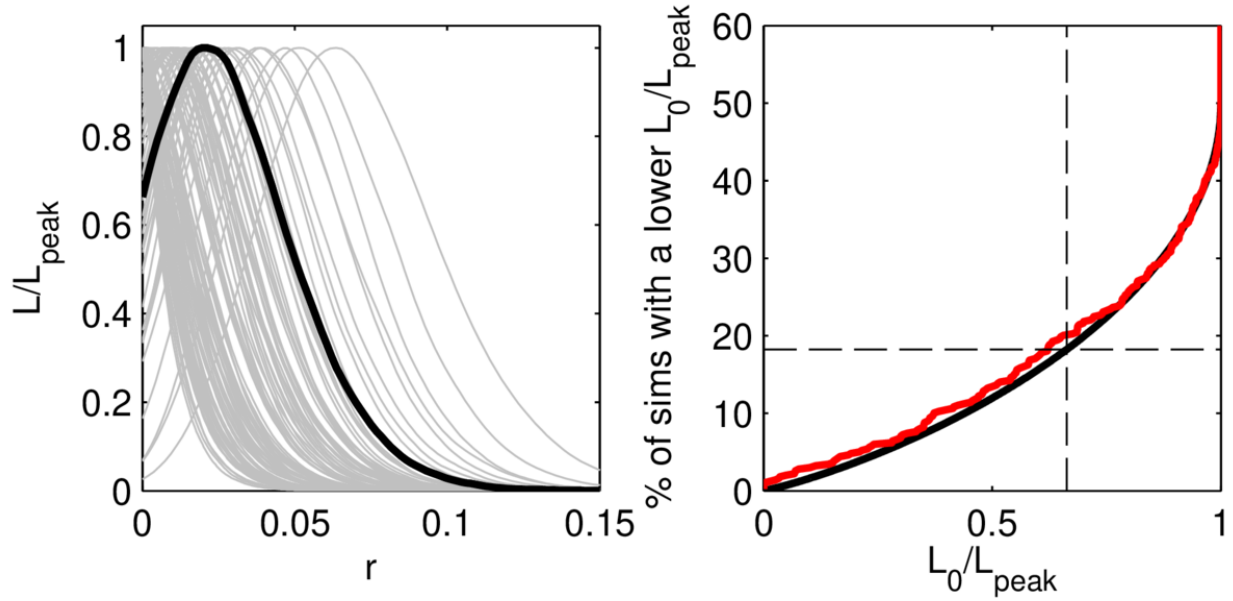
Figure 5.9: Zero-to-peak results from the BK15 analysis. Left: r likelihoods normalized by their maximum likelihood values for the region $r \geq 0$, where grey lines show simulations and the black line is the real data result. Right: CDF of the zero-to-peak ratio (value of the likelihood at $r = 0$ divided by it's maximum value) in red, and the simple analytic ansatz $\frac{1}{2}(1 - f(-2logL_0/L_{peak}))$ plotted in black, where f is the χ^2 CDF for one degree of freedom.

the wide prior curve in the center is $r = -0.022^{+0.042}_{-0.027}$, and for the narrow prior centered at $\beta_d = 1.53$ on the right is $r = -0.018^{+0.036}_{-0.028}$ as compared with the interval quoted in the BK15 paper which is $r = 0.02^{+0.021}_{-0.018}$.

5.5   95% Upper Limit

Another interval we can calculate is the 95% upper limit, which tells us at 95% confidence that the value of r lies below some value. For this calculation we consider only the positive region of our r likelihood curve, and integrate from zero out to the value of r for which 95% of the curve is contained. We calculate this interval for each simulation and histogram the results, shown in figure 5.11 with the narrow β_d prior results in the center, and the wide prior results on the right. The red line on each is the BK15 95% upper limit at $r = 0.07$, and the black line is the upper limit we calculate from real data. The blue dashed line is the upper
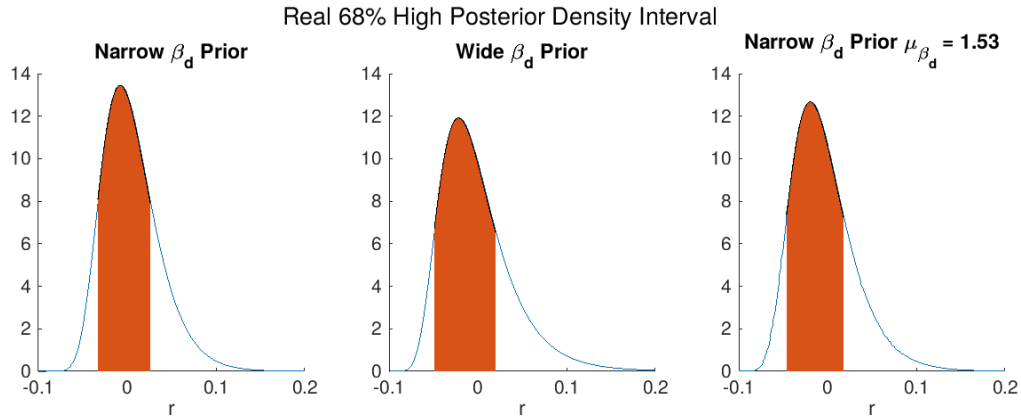
Figure 5.10: r posteriors calculated from the ILC analysis, where the curves are the likelihood curves normalized to integrate to 1. The orange shaded region shows the region of the 68% Highest Posterior Density Interval. The left panel shows results calculated with a narrow β_d prior, and the right side shows results calculated with a wide β_d prior.

limit for the data calculated with a narrow β_d prior centered at $\beta_d = 1.53$. We see that the distribution is centered at a higher value for the wide prior results versus the narrow prior, which is unsurprising considering the tales of these likelihood curves are longer as seen in figure 5.3. In the left panel are plotted the r posterior curves for the real data calculated with the different β_d priors, where the area under the likelihood curves is normalized to 1 for $r > 0$. The results for the real data put through our ILC analysis are quite close between different prior calculations, again with the wide prior result being slightly higher. The shapes of the normalized priors in the left hand plot are quite close, and we can see why our 95% upper limit calculations are close regardless of the β_d prior chosen. We find for both narrow β_d prior options that $r < 0.1$ at 95% confidence, and for the wide β_d prior that $r < 0.11$ at 95% confidence.

5.6   Conclusions

The results presented here are based off of a study on the BICEP/Keck 2015 data release. The BK15 analysis found a constraint $r < 0.07$ at 95% confidence, and after running a maximum likelihood search found unbiased results and $\sigma(r) = 0.020$. Our ILC analysis
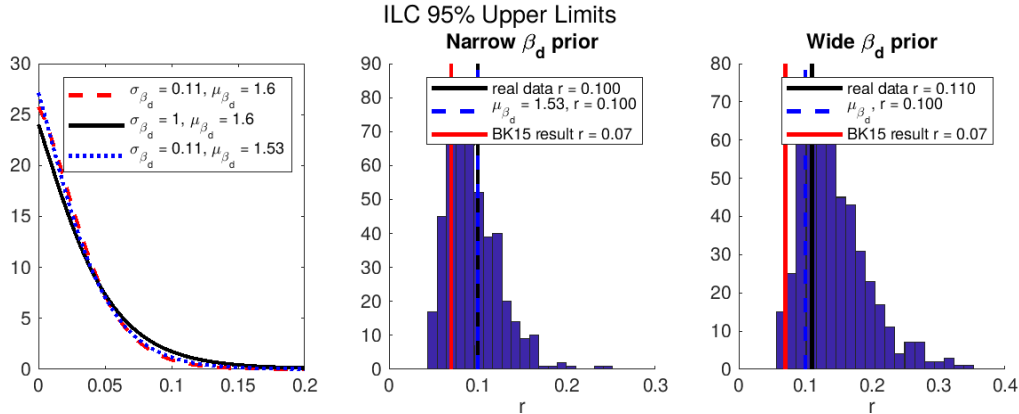
Figure 5.11: Distribution of the 95% upper limits of the r posteriors, where the region considered is $r \geq 0$. The left plot shows the posteriors for the data where the various β_d priors have been used. The center plot shows results for a narrow β_d prior, and the right shows results for a wide β_d prior. The blue dotted line in each plot shows the real data result for a narrow β_d prior centered at $\beta_d = 1.53$. The distributions contain results for the 499 simulated posterior curves, the red line is the BK15 95% upper limit calculated from real data, and the black line is the ILC analysis result from real data.

performed on the same data found the reasonable, yet looser constraint $r < 0.1$ at 95% confidence, and a maximum likelihood search found unbiased results and in the worst case $\sigma(r) = 0.031$.

The ILC analysis takes a much different approach to the data than the BK15 standard approach, making simple assumptions about the foreground content in the sky maps and marginalizing over the foreground parameter space in order to end up with a one dimensional r likelihood calculation, as opposed to the multiparameter search done in BK15. While our results are a bit less constraining than those of the BK15 analysis, their existence within the same magnitude of the BK15 results is quite promising.

The methods detailed in this paper could be modified to be able to support the data from the BICEP/Keck 2018 data release, which has data from the BICEP3 telescope in addition to the BICEP2 and Keck Array data. The BICEP3 telescope covers a larger area of sky than the BICEP2 style receivers, complicating the combination of maps performed in the ILC. In addition to the continued observations of the current BICEP3 and BICEP/Array telescopes, the CMB community is actively planning the CMB-S4 experiment, a Stage 4 ground-based

experiment which will collect more data from the CMB for the purposes of studying inflation, dark matter, and a host of other exciting fields. A map based ILC analysis alternative is being explored with the goal of performing the analysis on CMB-S4 data.

Bibliography

[1]   Scott Dodelson. *Modern Cosmology*. Academic Press, Elsevier Science, 2003.

[2]   Albert Einstein. "Die Feldgleichungen der Gravitation". In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften Berlin* (Jan. 1915), pp. 844–847.

[3]   A. Friedmann. "On the Curvature of Space". In: *General Relativity and Gravitation* 31 (Dec. 1999), p. 1991. DOI: `10.1023/A:1026751225741`.

[4]   Silvia Galli et al. "CMB constraints on dark matter models with large annihilation cross section". In: *Physical Review D* 80.2 (July 2009). DOI: `10.1103/physrevd.80.023505`. URL: `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.80.023505`.

[5]   Dorothea Samtleben, Suzanne Staggs, and Bruce Winstein. *The Cosmic Microwave Background for Pedestrians: A Review for Particle and Nuclear Physicists*. 2008. arXiv: `0803.0834 [astro-ph]`.

[6]   N. Aghanim et al. "Planck 2018 results. VI. Cosmological Parameters." In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. DOI: `10.1051/0004-6361/201833910`. URL: `https://www.aanda.org/articles/aa/full_html/2020/09/aa33910-18/aa33910-18.html`.

[7]   Alan H. Guth. "Inflationary universe: A possible solution to the horizon and flatness problems". In: *Phys. Rev. D* 23 (2 Jan. 1981), pp. 347–356. DOI: `10.1103/PhysRevD.23.347`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.23.347`.

[8]   Marc Kamionkowski and Ely D. Kovetz. "The Quest for B Modes from Inflationary Gravitational Waves". In: *Annual Review of Astronomy and Astrophysics* 54.1 (Sept. 2016), pp. 227–269. DOI: `10.1146/annurev-astro-081915-023433`. URL: `https://doi.org/10.1146/annurev-astro-081915-023433`.

[9]   George Lazarides. *Introduction to Inflationary Cosmology.* 2002. arXiv: `hep-ph/0204294 [hep-ph]`.

[10]   D. J. Fixsen et al. "The Cosmic Microwave Background Spectrum from the Full *COBE** FIRAS Data Set". In: *The Astrophysical Journal* 473.2 (Dec. 1996), pp. 576–587. DOI: `10.1086/178173`. URL: `https://iopscience.iop.org/article/10.1086/178173`.

[11]   N. Aghanim et al. "Planck 2018 results. I. Overview and the cosmological legacy of Planck". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A1. DOI: `10.1051/0004-6361/201833880`. URL: `https://www.aanda.org/articles/aa/full_html/2020/09/aa33880-18/aa33880-18.html`.

[12]   Matias Zaldarriaga and Uroš Seljak. "All-sky analysis of polarization in the microwave background". In: *Physical Review D* 55.4 (Feb. 1997), pp. 1830–1840. DOI: `10.1103/physrevd.55.1830`. URL: `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.55.1830`.

[13]   Maria Chiara Guzzetti et al. *Gravitational waves from inflation.* 2016. arXiv: `1605.01615 [astro-ph.CO]`.

[14]   Clive Dickinson. *CMB foregrounds - A brief review.* 2016. arXiv: `1606.03606 [astro-ph.CO]`.

[15]   R. Adam et al. "Planck intermediate results. XXX." In: *Astronomy Astrophysics* 586.A133 (2016). DOI: `https://doi.org/10.1051/0004-6361/201425034`. URL: `https://www.aanda.org/articles/aa/full_html/2016/02/aa25034-14/aa25034-14.html`.

[16]   A. Lazarian. "Tracing magnetic fields with aligned grains". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 106.1-3 (July 2007), pp. 225–256. DOI: `10.1016/j.jqsrt.2007.01.038`. URL: `https://doi.org/10.1016/j.jqsrt.2007.01.038`.

[17]   R. Adam et al. "Planck 2015 results". In: *Astronomy &amp Astrophysics* 594 (Sept. 2016), A1. DOI: `10.1051/0004-6361/201527101`. URL: `https://www.aanda.org/articles/aa/full_html/2016/10/aa27101-15/aa27101-15.html`.

[18]   M.-A. Miville-Deschênes et al. "Separation of anomalous and synchrotron emissions using WMAP polarization data". In: *Astronomy &amp Astrophysics* 490.3 (Sept. 2008), pp. 1093–1102. DOI: `10.1051/0004-6361:200809484`. URL: `https://www.aanda.org/articles/aa/abs/2008/42/aa09484-08/aa09484-08.html`.

[19]   U. Fuskeland et al. "SPATIAL VARIATIONS IN THE SPECTRAL INDEX OF POLARIZED SYNCHROTRON EMISSION IN THE 9 yr WMAP SKY MAPS". In: *The Astrophysical Journal* 790.2 (July 2014), p. 104. DOI: `10.1088/0004-637x/790/2/104`. URL: `https://doi.org/10.1088/0004-637X/790/2/104`.

[20]   A. Manzotti et al. "CMB Polarization *B*-mode Delensing with SPTpol and *Herschel*". In: *The Astrophysical Journal* 846.1 (Aug. 2017), p. 45. DOI: `10.3847/1538-4357/aa82bb`. URL: `https://doi.org/10.3847/1538-4357/aa82bb`.

[21]   P. A. R. Ade et al. "A demonstration of improved constraints on primordial gravitational waves with delensing". In: *Physical Review D* 103.2 (Jan. 2021). DOI: `10.1103/physrevd.103.022004`. URL: `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.103.022004`.

[22]   D. Barkats et al. "DEGREE-SCALE COSMIC MICROWAVE BACKGROUND POLARIZATION MEASUREMENTS FROM THREE YEARS OF BICEP1 DATA". In: *The Astrophysical Journal* 783.2 (Feb. 2014), p. 67. DOI: `10.1088/0004-637X/783/2/67`. URL: `https://iopscience.iop.org/article/10.1088/0004-637X/783/2/67`.

[23] Ruth Durrer. *The Cosmic Microwave Background: The history of its experimental investigation and its significance for cosmology.* 2015. arXiv: `1506.01907` [`astro-ph.CO`].

[24] P. A. R. Ade et al. "Improved Constraints on Primordial Gravitational Waves using Planck, WMAP, and BICEP/Keck Observations through the 2018 Observing Season". In: *Physical Review Letters* 127.15 (Oct. 2021). DOI: `10.1103/physrevlett.127.151301`. URL: `https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.127.151301`.

[25] Howard Hui and et al. *BICEP Array: a multi-frequency degree-scale CMB polarimeter.* 2018. arXiv: `1808.00568` [`astro-ph.IM`].

[26] P. A. R. Ade et al. "BICEP2. II. EXPERIMENT AND THREE-YEAR DATA SET". In: *The Astrophysical Journal* 792.1 (Aug. 2014), p. 62. DOI: `10.1088/0004-637x/792/1/62`. URL: `https://iopscience.iop.org/article/10.1088/0004-637X/792/1/62`.

[27] K. D. Irwin et al. "Time-division SQUID multiplexers". In: *Low Temperature Detectors.* Ed. by F. Scott Porter et al. Vol. 605. American Institute of Physics Conference Series. Feb. 2002, pp. 301–304. DOI: `10.1063/1.1457650`.

[28] E. S. Battistelli et al. "Functional Description of Read-out Electronics for Time-Domain Multiplexed Bolometers for Millimeter and Sub-millimeter Astronomy". In: *Journal of Low Temperature Physics* 151.3-4 (May 2008), pp. 908–914. DOI: `10.1007/s10909-008-9772-z`.

[29] P. A. R. Ade et al. "BICEP2/KECK ARRAY. VII. MATRIX BASED E/B SEPARATION APPLIED TO BICEP2 AND THE KECK ARRAY". In: *The Astrophysical Journal* 825.1 (June 2016), p. 66. DOI: `10.3847/0004-637x/825/1/66`. URL: `https://iopscience.iop.org/article/10.3847/0004-637X/825/1/66`.

[30] P. A. R. Ade et al. "Bicep2. III. INSTRUMENTAL SYSTEMATICS". In: *The Astrophysical Journal* 814.2 (Nov. 2015), p. 110. DOI: `10.1088/0004-637x/814/2/110`. URL: `https://iopscience.iop.org/article/10.1088/0004-637X/814/2/110`.

[31] C. L. Bennett et al. "NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP) OBSERVATIONS: FINAL MAPS AND RESULTS". In: *The Astrophysical Journal Supplement Series* 208.2 (Sept. 2013), p. 20. DOI: `10.1088/0067-0049/208/2/20`. URL: `https://iopscience.iop.org/article/10.1088/0067-0049/208/2/20`.

[32] Keir K. Rogers et al. "SILC: a new Planck internal linear combination CMB temperature map using directional wavelets". In: *Monthly Notices of the Royal Astronomical Society* 460.3 (May 2016), pp. 3014–3028. DOI: `10.1093/mnras/stw1121`. URL: `https://academic.oup.com/mnras/article/460/3/3014/2609411`.

[33] G. Hurier, J. F. Macı as-Pérez, and S. Hildebrandt. "MILCA, a modified internal linear combination algorithm to extract astrophysical emissions from multifrequency sky maps". In: *Astronomy &amp Astrophysics* 558 (Oct. 2013), A118. DOI: `10.1051/0004-6361/201321891`. URL: `https://www.aanda.org/articles/aa/full_html/2013/10/aa21891-13/aa21891-13.html`.

[34] P. A. R. Ade et al. "Constraints on Primordial Gravitational Waves Using Planck, WMAP, and BICEP/Keck Observations through the 2015 Season". In: *Physical Review Letters* 121.22 (Nov. 2018). DOI: `10.1103/physrevlett.121.221301`. URL: `https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.121.221301`.

[35] J. Delabrouille and J. -F. Cardoso. *Diffuse source separation in CMB observations.* 2007. arXiv: `astro-ph/0702198 [astro-ph]`.

[36] Rajib Saha et al. "CMB anisotropy power spectrum using linear combinations of WMAP maps". In: *prd* 78.2, 023003 (July 2008), p. 023003. DOI: `10.1103/PhysRevD.78.023003`. arXiv: `0706.3567 [astro-ph]`.

[37] Carl D. Meyer. "Generalized Inversion of Modified Matrices". In: *SIAM J. Appl. Math.* 24.3 (May 1973), pp. 315–323. DOI: `10.1137/0124033`.

[38] Samira Hamimeche and Antony Lewis. "Likelihood analysis of CMB temperature and polarization power spectra". In: *Physical Review D* 77.10 (May 2008). DOI: `10.1103/`

physrevd.77.103013. URL: `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.77.103013`.

[39] B. Thorne et al. "The Python Sky Model: software for simulating the Galactic microwave sky". In: *Monthly Notices of the Royal Astronomical Society* 469.3 (May 2017), pp. 2821–2833. DOI: `10.1093/mnras/stx949`. URL: `https://doi.org/10.1093/mnras/stx949`.

[40] and P. A. R. Ade et al. "Planck intermediate results. XXII. Frequency dependence of thermal emission from Galactic dust in intensity and polarization". In: *Astronomy & Astrophysics* 576 (Apr. 2015), A107. DOI: `10.1051/0004-6361/201424088`. URL: `https://www.aanda.org/articles/aa/full_html/2015/04/aa24088-14/aa24088-14.html`.

[41] Aaron M. Meisner and Douglas P. Finkbeiner. "MODELING THERMAL DUST EMISSION WITH TWO COMPONENTS: APPLICATION TO THE PLANCK HIGH FREQUENCY INSTRUMENT MAPS". In: *The Astrophysical Journal* 798.2 (Dec. 2014), p. 88. DOI: `10.1088/0004-637x/798/2/88`. URL: `https://iopscience.iop.org/article/10.1088/0004-637X/798/2/88`.

[42] Brian Hensley. "On the Nature of Interstellar Grains". Available at `http://arks.princeton.edu/ark:/88435/dsp0170795b03v`. PhD thesis. Princeton University, 2015.

[43] A G Kritsuk, S D Ustyugov, and M L Norman. "The structure and statistics of interstellar turbulence". In: *New Journal of Physics* 19.6 (June 2017), p. 065003. DOI: `10.1088/1367-2630/aa7156`. URL: `https://iopscience.iop.org/article/10.1088/1367-2630/aa7156`.

[44] Colin Bischoff. "Update to ML search results for alternate dust sims". BK internal memo. Dec. 2017.

[45]   and Y. Akrami et al. "iPlanck/i2018 results". In: *Astronomy &amp Astrophysics* 641
       (Sept. 2020), A11. DOI: 10.1051/0004-6361/201832618. URL: https://www.aanda.
       org/articles/aa/full_html/2020/09/aa32618-18/aa32618-18.html.