Testing Inflationary Cosmology with the BICEP1 and BICEP2 Experiments

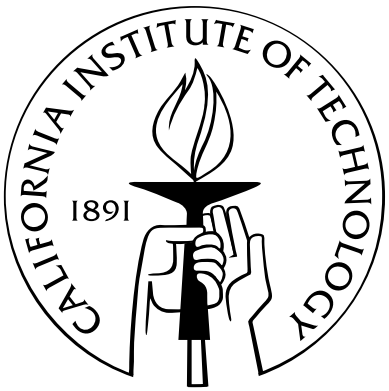
Thesis by

Randol Wallace Aikin

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended May 24, 2013)

Dedicated in loving memory to my brother Aaron and his passion for science.

Acknowledgments

First and foremost, I would like to acknowledge and thank the community of collaborators, advisors, post-docs, and students at Caltech who have made this pursuit such a rewarding academic experience. In particular, I thank my advisor, Prof. Sunil Golwala for a wonderful and productive academic relationship. I have enormously benefited from his rigor, generosity, and patience. I thank Prof. Jamie Bock for his tutelage and guidance, particularly during my early career at Caltech. I also owe a great deal to Prof. Andrew Lange, who we lost much to early, for his exemplary leadership and vision. I thank the outstanding post-docs and research scientists at Caltech and JPL from whom I have gained most of my experience and knowledge: Walt Ogburn, Jeff Filippini, Marcus Runyan, Jack Sayers, Martin Lueker, Zak Staniszewski, Bill Jones, Angiola Orlando, Darren Dowell, Hien Nguyen, Phil Wilson, and Rashmi Sudiwala. This group of scientists provided the core of my education at Caltech. In particular, I would like to acknowledge Walt Ogburn and Jeff Filippini as invaluable resources for my graduate school education. I thank them for their generosity with their time and tremendous talents. I would like to thank all of my fellow Caltech graduate students, specifically Justus Brevik and Zeeshan Ahmed for their camaraderie, and from whom I learned a great deal.

I would further like to thank and acknowledge the BICEP2 collaboration, beginning with the project's principal investigator, John Kovac. Much of my graduate school education has been the process of downloading John Kovac's brain into mine. I would like to thank him for, among other things, high bandwidth. I am extremely grateful for having had the opportunity to learn from John's tenacity, critical thinking, and methodology. I would like to thank our collaborators at Harvard, Stanford, University of Minnesota, and UCSD. Specifically, Colin Bischoff, Clem Pryke, Chris Sheehy, and Jamie Tolan, who contributed substantially to much of the analysis content of this thesis. I also thank Steffen Richter, BICEP2's three-time South Pole winter-over, for his dedication, professionalism, and mastery of the BICEP2 telescope. It is safe to say that BICEP2's success would not have been possible without Steffen.

I would like to thank the BICEP1 collaboration, on whose work I relied heavily. The analyses by Cynthia Chiang and Yuki Takahashi served as blueprints for the three-year analysis. I thank them, and the broader BICEP1 team for sharing guidance, insight, and experience. I also thank Denis

Barkats, who led the three-year analysis, for his leadership and dedication to the task.

The construction and operation of BICEP2 would not have been possible without the tremendous support provided by the wonderful South Pole Station support staff. I would also like to thank the staff at Caltech for making the project possible. I would particularly like to thank Kathy Deniston for keeping the ship afloat.

I am extremely lucky and grateful to have the tremendous support of my wonderful family. I am eternally grateful for their sustaining support, confidence, and encouragement. Dad, Carter, Diane, Taylor, Purvi, Terry, Juleigh, Niki, Ed, Terry, Carl, and everyone else—thank you. I would also like to thank my incredible community of friends in Los Angeles for keeping me sane and grounded through graduate school. I won't name names. You know who you are.

Abstract

Recent observations of the temperature anisotropies of the cosmic microwave background (CMB) favor an inflationary paradigm in which the scale factor of the universe inflated by many orders of magnitude at some very early time. Such a scenario would produce the observed large-scale isotropy and homogeneity of the universe, as well as the scale-invariant perturbations responsible for the observed (10 parts per million) anisotropies in the CMB. An inflationary epoch is also theorized to produce a background of gravitational waves (or tensor perturbations), the effects of which can be observed in the polarization of the CMB. The E-mode (or parity even) polarization of the CMB, which is produced by scalar perturbations, has now been measured with high significance. Contrastingly, today the B-mode (or parity odd) polarization, which is sourced by tensor perturbations, has yet to be observed. A detection of the B-mode polarization of the CMB would provide strong evidence for an inflationary epoch early in the universe's history.

In this work, we explore experimental techniques and analysis methods used to probe the B-mode polarization of the CMB. These experimental techniques have been used to build the BICEP2 telescope, which was deployed to the South Pole in 2009. After three years of observations, BICEP2 has acquired one of the deepest observations of the degree-scale polarization of the CMB to date. Similarly, this work describes analysis methods developed for the BICEP1 three-year data analysis, which includes the full data set acquired by BICEP1. This analysis has produced the tightest constraint on the B-mode polarization of the CMB to date, corresponding to a tensor-to-scalar ratio estimate of $r = 0.04 \pm 0.32$, or a Bayesian 95% credible interval of $r < 0.70$. These analysis methods, in addition to producing this new constraint, are directly applicable to future analyses of BICEP2 data. Taken together, the experimental techniques and analysis methods described herein promise to open a new observational window into the inflationary epoch and the initial conditions of our universe.

Contents

List of Figures

List of Tables

The BICEP2 Collaboration

California Institute of Technology/Jet Propulsion Laboratory

Randol Aikin, Jamie Bock, Tony Bonetti, Darren Dowell, Jeff Filippini, Sunil Golwala, Viktor Hristov, Andrew Lange (*In Memoriam)*, Hien Nguyen, Zak Staniszewski, Grant Teply, Anthony Turner, Phil Wilson

Cardiff University

Peter Ade, Rashmi Sudiwala

Case Western Reserve University

John Ruhl

CEA Grenoble

Lionel Duband

Harvard University

Colin Bischoff, John Kovac, Chin Lin Wong

National Institute of Standards and Technology

Justus Brevik, Kent Irwin

Stanford University

Sarah Stokes Kernasovskiy, Chao-Lin Kuo, R. Walt Ogburn, Jamie Tolan

University of British Columbia

Mark Halpern

University of California, San Diego

Jon Kaufman, Brian Keating, Angiola Orlando

University of Minnesota

Stefan Fliescher, Clem Pryke, Chris Sheehy

University of Toronto

Steve Benton, Barth Netterfield

South Pole Winterover

Steffen Richter

Chapter 1

Open questions in modern cosmology

Over the course of the last century, our view of the universe has been revolutionized. Once believed to be eternal, static, and infinite, we now understand our universe to be dynamic, expanding, and finite, governed by universal physical laws and described by a limited number of physical constants and cosmological parameters. This revolution has been made possible by a confluence of laboratory experiments, observations of the night sky, and the refinement of physical theory.

This effort to describe the history, evolution, and ultimate fate of the universe is by no means complete. We continue to push our physical understanding of the universe back to the beginning of time itself, requiring new experiments, new theory, and perhaps new physics. In this work, we present new experimental techniques, analysis methods, and observations to address the following open questions in modern cosmology:

- What dominated the energy density of the very early universe?
- What is the source of the density perturbations responsible for structure in our universe?
- How did the universe come to be nearly perfectly geometrically flat?
- How did causally disconnected regions of space come to be in apparent thermal equilibrium?

1.1 ΛCDM Cosmology

Experimental and theoretical efforts over the last century have culminated in a widely accepted model of the early universe, referred to as ΛCDM cosmology (also called standard, or concordance cosmology). ΛCDM cosmology can be broadly summarized as follows: i) The universe's geometry is described by the Friedmann-Robertson-Walker metric. ii) The universe principally consists of radiation, baryonic matter, cold dark matter (CDM), and so-called "dark energy" (Λ). iii) The universe is expanding according to the Friedmann equations, and that expansion is today being driven by dark energy. iv) The universe began in an extremely high energy, high density state and has

since been expanding and cooling over the last 13.7 billion years. v) The universe's history consists of three epochs: radiation domination, matter domination, and dark energy domination, and, perhaps at very early times, an inflationary epoch. In this section, we will describe the cornerstones of ΛCDM cosmology. In describing ΛCDM in detail, we rely heavily on three sources: "Cosmological Inflation and Large-Scale Structure" (Little and Lyth 2000), "Modern Cosmology" (Dodelson 2003), and Daniel Baumann's lecture notes, "The Physics of Inflation."[1] As in these texts, units are chosen such that $c = 1$, only entering c into expressions when a change of unit is to be emphasized.

1.1.1 The Friedmann-Robertson-Walker Universe

One of the first, and most critical, assertions of modern cosmology is that the geometry is described by the Friedmann-Robertson-Walker (FRW) metric, with a line element given by:

$$ds^2 = -dt^2 + a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right], \tag{1.1}$$

where $a(t)$ is the scale factor , k describes the geometric curvature , $d\Omega$ is a volume element given by $d\Omega = d\theta^2 + \sin^2\theta d\phi$, and t is the physical time. The expansion of the universe is thus described by the time evolution of $a(t)$. We can perform a useful change of variable to conformal time, defined as:

$$\tau = \int \frac{dt}{a}. \tag{1.2}$$

Recasting Equation 1.1 in conformal time, we find the line element to be:

$$ds^2 = a(\tau)^2 \left[-d\tau^2 + \frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right]. \tag{1.3}$$

Given two points in a universe with a time-evolving scale factor, we can write down a conformal time interval, $d\tau$. If the distance that separates these two points is less than $c\,d\tau$, these points are said to be in causal contact, that is, within their respective light cones. If the distance that separates these points exceeds $c\,d\tau$, these points are not in causal contact. This is the power of working in terms of conformal time: it immediately allows us to establish whether two points can be in causal contact. Put another way, working in terms of conformal time, null geodesics (light cones) are drawn as straight lines, regardless of the time evolution of the scale factor. Setting $ds = 0$, the null geodesic is simply:

$$\frac{dr}{d\tau} = \pm 1. \tag{1.4}$$

[1] `http://www.damtp.cam.ac.uk/user/db275/TEACHING/INFLATION/Lectures.pdf`

In the spirit of the cosmology texts referenced above, we refer to any finite τ as a "cosmological horizon."

1.1.2 Friedmann Equations

The first and second Friedmann equations, introduced by Alexander Friedmann in 1922, together relate the time evolution of the scale factor to the content and the geometric curvature of the universe. The first and second Friedmann equations are as follows (letting the Planck mass $M_{pl} = (8\pi G)^{-1/2}$):

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{\rho}{3M_{pl}^2} - \frac{k}{a^2} \tag{1.5}$$

$$\frac{\ddot{a}}{a} = -\frac{1}{6M_{pl}^2}(\rho + 3p). \tag{1.6}$$

Here ρ is the energy density, p is the pressure, k is again the geometric curvature, and derivatives are taken with respect to physical time t. In this expression, any cosmological constant Λ is folded into the total energy density, ρ.

The interpretation of Friedmann equations is simple, yet startling: the expansion of the universe is driven by the equation of state of the dominant form of energy density of the universe. The time dependence of ρ is dictated by the continuity relation:

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p) = -3H(\rho + p). \tag{1.7}$$

The geometric curvature term k, now measured to be very close to zero, additionally dictates the time evolution of a. With the Friedmann equations in hand, we can examine the known history of the universe.

1.1.3 Epochs

The universe can be roughly divided into three semi-distinct epochs: radiation domination, matter domination, and dark energy domination. During these epochs, the time evolution of the scale factor changes as a result of the corresponding equations of state. The equation of state is given by:

$$\rho = \rho_0 a^{-3(w+1)}, \tag{1.8}$$

where $w = 0$ for matter, $w = 1/3$ for radiation, and $w = -1$ for dark energy, and ρ_0 is the energy density of the current epoch. Assuming a "single-component" universe, we can find reasonable approximations for the time evolution of the scale factor during radiation domination, matter

domination, and dark energy domination:

$$a_\gamma \propto t^{1/2} \tag{1.9}$$

$$a_{\rm M} \propto t^{2/3} \tag{1.10}$$

$$a_\Lambda \propto \exp\left(\sqrt{\frac{\Lambda}{3}}t\right). \tag{1.11}$$

At very early times, radiation dominated the energy density of the universe. As the universe cooled and expanded, the radiation energy density dropped relative to the matter density, and by a redshift of $z \sim 3000$, the universe reached matter-radiation equality. As expansion continued, the relative density of radiation to matter continued to decrease as a^{-1}, giving rise to the epoch of matter domination. While the matter density has diluted with the expansion of the universe, the dark energy density has remained constant. For this reason, dark energy is sometimes referred to as vacuum energy. Current observations suggest that today roughly 70% of our universe consists of this dark energy. The remaining 30%, consisting almost entirely of matter, plays an important role in the time evolution of a. Today, the radiation density is negligible compared to these other species.

As we extrapolate backwards in time, we might naively expect that radiation domination continues as we asymptotically approach $t = 0$. This is not unreasonable: matter and dark energy must certainly be subdominant. However, assuming radiation domination back to $t = 0$ presents two conflicts: the horizon problem and the flatness problem, which we describe below.

1.1.4 Horizon problem

The horizon problem addresses the question, how did distant regions of the universe come to be within causal contact in the past? For illustrative purposes, we will examine the horizon problem at recombination; a distinct period of time in which protons and electrons combined for the first time to form neutral hydrogen. Current observations have measured recombination to be at $z \sim 1100$ and $t = 380,000$ years. As before, we can calculate the causal horizon:

$$\tau = \int \frac{dt}{a} = \int \frac{da}{a^2 H}. \tag{1.12}$$

This is often re-expressed in terms of the co-moving Hubble radius, $(aH)^{-1}$:

$$\tau = \int (aH)^{-1} d\ln a. \tag{1.13}$$

Making use of Friedmann's equations and ignoring curvature, this can be written as:

$$\tau \propto \int a^{-1/2+3/2w} da. \tag{1.14}$$

We consider a few distinct cases. To begin, we calculate the causal horizon at the time of recombination and assume radiation domination. In this case, $\tau_{\text{recomb}} \sim 0.2$ Mpc. To put this into perspective, this is a factor of a few smaller than the current distance between the Milky Way and the Andromeda galaxy. This presents a distinct conflict: we observe the universe to be in apparent thermal equilibrium over scales much larger than τ_{recomb}, suggesting that at some point in the early universe's history, the causal horizon must have been much larger than τ_{recomb}. While we have assumed radiation domination, the same conclusion is reached when matter domination is assumed. Our naive guess is therefore problematic. As we extrapolate back to $t = 0$, assuming radiation or matter domination forces us to conclude that distant regions of the universe have never been in causal contact. We thus find it plausible that our initial guess about the very early universe was wrong.

It is useful to illustrate the horizon problem in a co-moving space-time diagram (Figure 1.1). Points displaced along the r axis are separated in space, while points displaced along the τ axis are separated in conformal time. Since we are working in co-moving coordinates, null geodesics are always diagonal, regardless of the evolution of $a(t)$ (Equation 1.4). "Light cones" represent the volume of space-time in which points are said to be in causal contact, that is, their separation in space is less than the light travel time between them. At the present epoch, the cosmological horizon is τ_0, and at recombination the cosmological horizon is τ_{recomb} (spanned by a corresponding cosmological horizon distance r_{recomb}). At τ_{recomb}, we can draw light cones corresponding to the causal horizons of two regions of space at opposite ends of our observable universe, separated by $2(r_0 - r_{\text{recomb}}) \approx 2r_0$. Today, we observe these two distant regions to be in apparent thermal equilibrium, but assuming radiation domination, they were never in causal contact. We will find later that postulating a different behavior of τ at early times gives us a possible solution to the horizon problem.

Considering the geometric curvature of the universe, we find further evidence that something else besides radiation or matter dominated the energy density of the early universe.

1.1.5 Flatness problem

Today, the geometric curvature of the universe, represented in Friedmann's equations by k, has been measured to be very close to flat. We can recast k as a curvature density, Ω_k, where $\Omega_k = k/a^2H^2$. In an over-dense universe, $\Omega_k > 0$, and the universe eventually re-collapses on itself. In an under-dense universe, $\Omega_k < 0$, and the universe diverges. Any perturbation from perfect flatness ($\Omega_k = 0$)

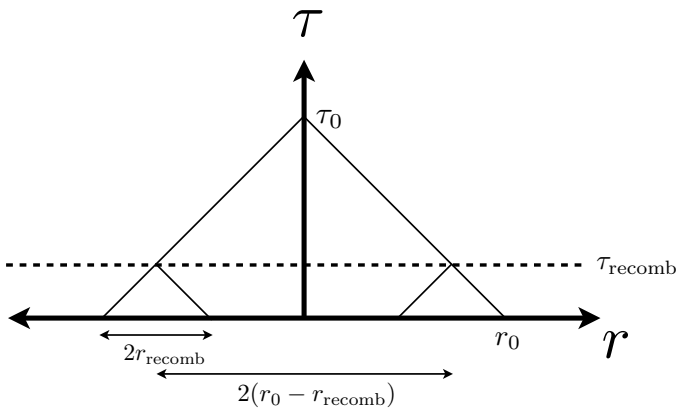


Figure 1.1: A graphical representation of the horizon problem. Diagonal lines represent null geodesics, or light cones, extending from the current epoch back to the beginning of physical time. Without inflation, the cosmological horizon today is $2r_0$. At the time of recombination, the causal horizon was $2r_{\rm recomb}$. The horizon problem is the tension between the observation that $r_{\rm recomb} \ll r_0$, and yet the universe is in thermal equilibrium over scales $\sim 2r_0$.

will quickly diverge into one of these scenarios. The tightest constraint on Ω_k comes from observations of the acoustic peaks of the temperature anisotropies of the cosmic microwave background (CMB, described in detail in Section 1.2). Observations from the WMAP satellite, together with constraints from the South Pole Telescope (SPT) and the Atacama Cosmology Telescope (ACT) have constrained $\Omega_k = 0.001 \pm 0.012$ (Hinshaw et al. 2012).

Because of the a^{-2} dependence of Ω_k, the fact that today we measure Ω_k to be close to zero implies that in the past it must have been much closer to zero. At the time of recombination, for instance, we find current constrains imply that $\Omega_k < 1 \times 10^{-8}$. How did the universe come to be so close to perfectly flat? This is the so-called "flatness problem."

1.2 The Cosmic Microwave Background

Perhaps the single most important tool for our current understanding of the content, dynamics, and history of the universe is the cosmic microwave background (CMB). The CMB was first discovered by Arno Penzias and Robert Wilson in 1965. Their observation was of an "effective zenith noise temperature of the 20-foot horn-reflector antenna [...] about 3.5 K higher than expected. This excess temperature is, within the limits of our observations, isotropic, unpolarized, and free from seasonal variations..." (Penzias and Wilson 1965).

The excess zenith temperature measured by Penzias and Wilson is the remnant radiation from the epoch of recombination. At $t < 380,000$ years, the baryonic content of the universe was entirely ionized. The universe was a hot plasma of primarily H and He, with small amounts of heavier

elements such as D, Be, and Li. Thompson scattering by free electrons efficiently processed photons, resulting in a mean free path of $\lambda = (n_e \sigma_e)^{-1}$, where n_e is the electron number density, and σ_e is the Thompson cross-section. As the universe expanded and cooled, electrons and protons combined to form H, dropping n_e considerably. Soon, the mean free path was comparable to the horizon scale and the photons were able free-stream across the universe. Today, we measure that photon background at a temperature of 3 K due to the expansion redshift since the time of emission.

While largely isotropic and homogeneous and nearly perfectly thermalized, small temperature fluctuations in the CMB have been revealed by a large number of telescopes, unveiling underlying anisotropy in the matter distribution in the early universe. These anisotropies have been a powerful tool for constraining cosmology. The characteristic angular size of the anisotropies have served as a standard ruler, while the depth of the gravitational potential wells have acted as a sensitive probe of the content of the early universe.

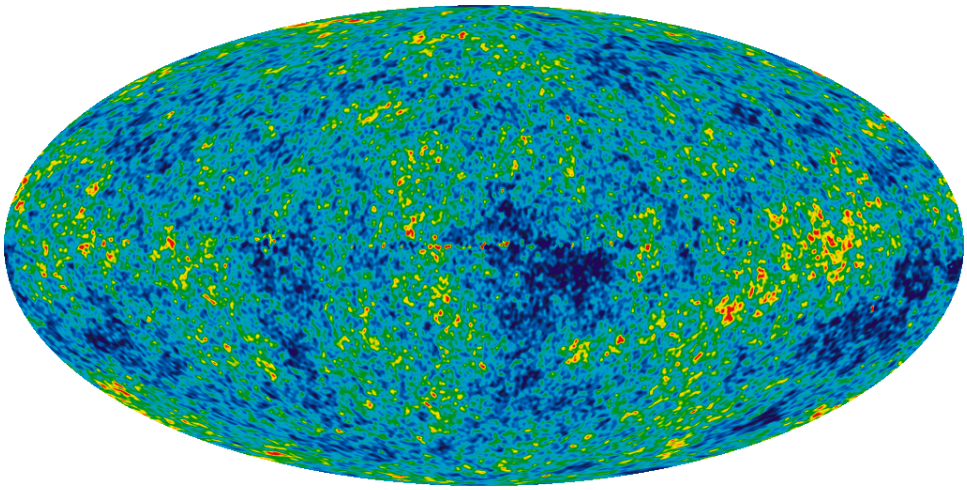


Figure 1.2: The Cosmic Microwave Background as seen by the WMAP satellite, courtesy of the WMAP science team (Hinshaw et al. 2012). After removing the mean temperature, the dipole due to our peculiar motion in the universe, and emission from our galaxy, primordial temperature anisotropies have been revealed at a level of 10 parts per million. Dark blue regions (counter-intuitively) represent over-dense regions, resulting in a temperature decrement, due to gravitational redshift as photons climb out of the dark matter potential wells. Warmer (red) regions represent under-dense regions, resulting in less gravitational redshift. The color stretch is ± 200 μK.

The CMB has been a remarkably powerful tool for constraining cosmological parameters for two reasons: first, its statistical properties can be described by straightforward semi-analytic equations that rely on a limited number of free parameters. This is due in part to the fact that at the time of recombination, the universe is well described by linear perturbation theory, before physical processes give rise to non-linear dynamics. Second, as a stochastic background, the CMB contains Gaussian-distributed independent modes, allowing a simple statistical analysis for cosmological parameter constraints. This will be described in detail in later sections.

1.2.1 Temperature anisotropies and angular power spectra

The discovery of this "excess temperature" initialized a broad effort to map the temperature anisotropies of the CMB to high precision. This effort has involved decades of research, three space satellites, dozens of ground-based experiments, and two Nobel prizes. The temperature anisotropy has now been mapped to exquisite precision by a number of experiments. The outcomes of these experiments are often presented as an angular power spectrum on the sphere, which we briefly review here.

We begin with the temperature anisotropy described as a function of position on the sky, $T(\hat{n})$, normalized by T_0. Here $\hat{n}$ is a position vector on the unit sphere. The temperature anisotropy field can be expanded into spherical harmonics as:

$$\frac{T(\hat{n})}{T_0} = 1 + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^T Y_{\ell m}(\hat{n}). \tag{1.15}$$

Here, the $Y_{\ell m}$ are the spherical harmonic basis functions and $a_{\ell m}^T$ are the coefficients of the expansion, given by:

$$a_{\ell m}^T = \int \frac{T(\hat{n})}{T_0} Y_{\ell m}^*(\hat{n}) d\hat{n}. \tag{1.16}$$

For most cosmological parameter estimation, we collapse these expansion coefficients over m. This is because the m modes are a coordinate- and position-dependent quantity. If we were to rotate the sky, or move to some other region of the observable universe, the "realization" of m-modes would be different, but, of course, we assume the cosmology to be the same[2]. Collapsing over m, we arrive at the canonical angular power spectrum:

$$C_\ell^{TT} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \left(a_{\ell m}^{T*} a_{\ell m}^T \right). \tag{1.17}$$

It is conventional to multiply the angular power spectrum by $\ell(\ell+1)/(2\pi)$:

$$\mathcal{C}_\ell^{TT} \equiv \frac{\ell(\ell+1)}{2\pi} C_\ell^{TT}. \tag{1.18}$$

In practice, there are a number of statistical and observational realities that restrict our ability to precisely measure C_ℓ. The derivation of these factors can be found in a number of references, including an excellent review, "CMB Anisotropies" by Hu and Dodelson 2002. As there are a finite number of m modes at each angular scale ℓ, there is a fundamental sample variance limit to the precision with which we can measure the power spectrum C_ℓ. This is the so-called "cosmic variance"

[2]While this is generally true, there are cosmological models that do invoke a position or coordinate dependence to cosmological parameters. To test these models, we must keep track of the m-modes.

limit:

$$\Delta C_\ell = \frac{2}{2\ell+1} C_\ell^2. \tag{1.19}$$

Furthermore, our ability to measure the power spectrum degrades as our sky coverage decreases, for the simple reason that we are sampling fewer modes. If we cover only a fraction of the sky, given by $f_{\rm sky}$, our uncertainty becomes:

$$\Delta C_\ell = \frac{2}{(2\ell+1) f_{\rm sky}} C_\ell^2. \tag{1.20}$$

Additionally, we inevitably have a noise from detectors, weather, amplifiers, etc. Given a noise power spectrum N_ℓ, we can write the resulting uncertainty as:

$$\Delta C_\ell = \frac{2}{(2\ell+1) f_{\rm sky}} (C_\ell + N_\ell)^2. \tag{1.21}$$

Finally, when making angular power spectra, we must combine the power into finitely sized bins[3] in ℓ. Given some bin width $\Delta\ell$, we find the resulting uncertainty to be:

$$\Delta C_\ell = \frac{2}{(2\ell+1) f_{\rm sky} \Delta\ell} (C_\ell + N_\ell)^2. \tag{1.22}$$

As experimentalists, we can draw several heuristic conclusions from the above. At the largest angular scales, sample variance dominates, occurring when ℓ is comparable to the observed field. At the smallest angular scales, N_ℓ dominates, since the noise power spectrum will increase at angular scales much smaller than the beam scale of the instrument.

1.2.2 Cosmological parameter constraints

Here we briefly describe how the CMB can be used as a probe of cosmological parameters. Once a major topic of theoretical research, models of the temperature anisotropy power spectrum are now highly refined and computationally efficient. Now theoretical power spectra can now be readily generated for any cosmology one chooses, using tools such as `CMBFAST`[4] and `CAMB`.[5] Hu and Dodelson 2002 note that the parameterization of the CMB power spectrum can be broadly separated into three categories: initial conditions, the energy content of the universe before recombination, and the energy content after recombination. I will use this distinction as well to describe the anatomy of the CMB

[3] In practice, because a given experiment has finite resolution in ℓ, this bin size is chosen such that finitely sampled bins are uncorrelated. In the limit that ℓ bins are highly correlated, Equation 1.22 no longer applies. We cannot therefore arbitrarily shrink $\Delta\ell$ to improve our statistical uncertainty. As usual in observational cosmology, there is no such thing as a free lunch.

[4] `http://lambda.gsfc.nasa.gov/toolbox/tb_cmbfast_ov.cfm`

[5] `http://camb.info/`

power spectrum.

The initial conditions of the CMB are perhaps the least understood. Describing these initial conditions remains a critical goal for modern experimental cosmology. While we might regard any number of physical constants as "initial conditions," we restrict ourselves to treating only the inflaton potential as an initial condition. As we will describe in Section 1.3, the inflaton potential manifests itself as the spectral index of the CMB power spectrum, and the B-mode polarization of the CMB.

The energy content of the universe prior to recombination is now relatively well understood. The acoustic peaks of the CMB power spectrum are sensitive probes of these parameters for the simple reason that the dynamics of the oscillations of the baryon-photon plasma in the early universe are largely governed by gravitational potential wells and photon pressure.

Finally, the angular scale of the first peak of the CMB acts as a "standard ruler", which is a known physical distance. This standard ruler can be used as a probe of the geometry of the universe, which in turn serves as a probe of the energy content of the universe after the epoch of recombination. Additionally, anisotropies from the epoch of recombination are later re-scattered by free electrons ionized by star light during the epoch of reionization. These anisotropies can be used as a probe for understanding when the universe became largely ionized.

To date, the CMB has been an extremely powerful tool for probing the energy content of the universe since very early times. The CMB promises more information still regarding the initial conditions of our universe, including the nature and energy scale of inflation.

1.3 Inflation

In his seminal 1980 paper "Inflationary Universe: A possible solution to the horizon and flatness problems," Alan Guth posited a solution to the problems outlined in previous sections: what if at some very early time in the universe's history, the causal horizon shrank? We take this to be the fundamental condition for inflation. Expressed more formally, we treat inflation as the scenario in which $d/dt((aH)^{-1}) < 0$. Considering the behavior of τ, this would mean that τ could assume an arbitrarily large negative value at very early times.

We consider one specific inflationary scenario: one in which the energy density of the universe is dominated by a vacuum energy with an equation of state $\rho = \rho_0 a^0$. Taking Equation 1.14 and setting $w = -1$, we can examine the behavior of τ as we roll the clock backwards from the end of inflation to the beginning of inflation. We find the behavior of τ we desire: τ can take on very large negative values at early times with inflationary expansion:

$$\tau \propto \int_{a_{\rm end}}^{a_{\rm start}} a^{-2} da \propto \frac{1}{a_{\rm end}} - \frac{1}{a_{\rm start}}. \tag{1.23}$$

In this way, inflation offers a possible solution to the horizon problem. This is illustrated in Figure 1.3. In the inflationary paradigm, τ does not reach a finite value as we extrapolate back to the beginning of physical time. Rather, the process of inflation shrinks the causal horizon by orders of magnitude, allowing τ to assume large negative values before inflation began.

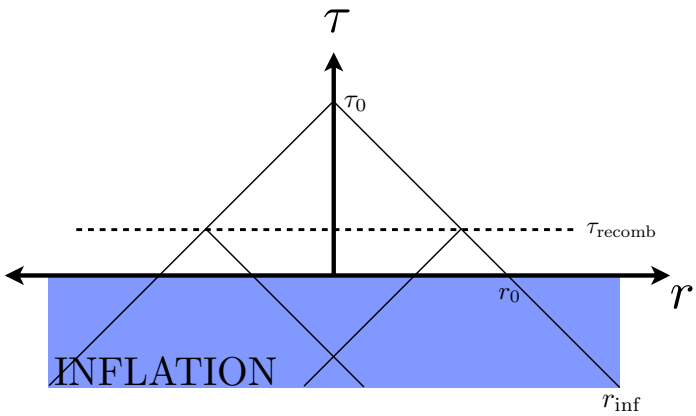


Figure 1.3: The inflationary paradigm's solution to the horizon problem. With inflation, causal time extends to very large negative values, stretching the cosmological horizon, $r_{\rm inf}$, to scales well beyond r_0. In this scenario, seemingly causally disconnected regions of space have the opportunity to reach thermal equilibrium.

Additionally, we can see how the process of inflation addresses the flatness problem. If we consider the Friedmann equation (1.5), as the scale factor a grows exponentially during inflation, any curvature that is present at the beginning will be exponentially diluted.

Perhaps the most compelling feature of inflation theory is that it provides a mechanism whereby scalar perturbations can be generated with a power spectrum that is today consistent with CMB observations. We explore this point in further detail in subsequent sections.

1.3.1 The inflationary potential

We have been able to address the flatness and horizon problems, but only by invoking some new as-of-yet undiscovered energy content of the universe that fully dominates the early universe. So far in our arguments, our only "evidence" for inflation theory is that it satisfies the conditions that it was invented to meet! Be that as it may, we can begin to broadly and vaguely define the characteristics of this new form of energy necessary to produce the dynamics required of an inflationary scenario.

Let us begin by assuming the existence of a scalar field, $\phi(\vec{x}, t)$. A scalar field is chosen solely because it is the simplest thing we can imagine. Like pressure, humidity or temperature, it is simply a scalar number that we can assign to any point in space and time. Next, let $V(\phi)$ be the potential energy density of the field (often referred to as the inflationary potential). If we assume that the

scalar field is homogenous in space, then we can write the energy density and pressure as:

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{1.24}$$

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{1.25}$$

We can substitute ρ and p into the continuity equation (Equation 1.7) to find the equation of motion:

$$\ddot{\phi} + 3H\dot{\phi} = -\frac{dV}{d\phi}. \tag{1.26}$$

In the event of a non-evolving scale factor ($H = 0$), the solution to Equation 1.26 will be oscillatory in time. The factor of H thus acts as a friction term, and is sometimes correspondingly referred to as Hubble friction. In the next section, we will examine the conditions for the scalar potential and its time evolution in order for inflation to occur.

1.3.2 Slow-roll inflation

We begin by recalling the condition for inflation: $d/dt((aH)^{-1}) < 0$, which is equivalent to the condition that $\ddot{a} < 0$. Examining the second Friedmann equation (Equation 1.6), we find that this condition is satisfied if the potential dominates over the kinetic term ($\dot{\phi}^2 < 2V$). If the second derivative term ($\ddot{\phi}$) is large, the kinetic term will quickly grow, and inflation will abruptly end. On the other hand, if the second derivative is zero and the initial condition for inflation is satisfied, then the potential term will always dominate, and inflation will never end. We therefore require a potential that is nearly flat (yielding a small first derivative), but has a non-zero second derivative. Slow-roll inflation describes an inflationary scenario that satisfies these criteria.

We parameterize these conditions as the slow-roll parameters:

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} \tag{1.27}$$

$$\eta = \frac{\dot{\epsilon}}{H\epsilon}. \tag{1.28}$$

The slow-roll conditions are that both of these parameters be much less than unity. Referring back to the Friedmann equations, this is equivalent to the approximations:

$$H^2 \simeq \frac{V(\phi)}{3M_{pl}^2} \tag{1.29}$$

$$3H\dot{\phi} \simeq -V'(\phi). \tag{1.30}$$

Let's consider both slow-roll conditions. Again, making use of the Friedmann equations, we can

see that $\varepsilon \ll 1$ is equivalent to the condition stated earlier:

$$\frac{3}{2}\left(1+\frac{p}{\rho}\right) = \frac{3}{2}\left(1+w\right) < 1. \tag{1.31}$$

Again, we find that something with with an equation of state $w < -1/3$ fits the bill. The second condition, $\eta \ll 1$, is simply requiring that ε remain less than one for a considerable period of (co-moving) time, so that the universe may expand by many orders of magnitude.

The slow-roll parameters can also be recast in terms of the inflationary potential and its derivatives with respect to ϕ. To recast ε in terms of V, we first consider Equations 1.29 and 1.30:

$$\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \simeq \frac{\dot{\phi}^2}{M_{pl}^4 H^2}. \tag{1.32}$$

Using the second Friedmann equation, we substitute $\dot{\phi}^2 = -3M_{pl}^2\dot{H}$. We therefore can define an alternate slow-roll parameter ϵ_V, which is equivalent to ε up to factors of order unity:

$$\epsilon_V \equiv M_{pl}^2\left(\frac{V'}{V}\right)^2. \tag{1.33}$$

Similarly for η, we begin by taking the time-derivative of Equation 1.30:

$$3\dot{H}\dot{\phi} + 3H\ddot{\phi} = V''(\phi)\dot{\phi}. \tag{1.34}$$

Performing similar substitutions as before, we arrive at the alternate second slow-roll parameter:

$$\eta_V \equiv M_{pl}^2\frac{V''(\phi)}{V(\phi)}. \tag{1.35}$$

One particular inflationary model to consider is illustrated in Figure 1.4, usually referred to as "Mexican hat" inflation (we will call it "sombrero" inflation). While "sombrero" inflation was originally invented as a mechanism for generating spontaneous symmetry breaking, we choose sombrero inflation as a vehicle for a) illustrating features common to all slow-roll potentials, and b) giving the reader a sense for the wide range of theoretical inflationary potentials that exist in the literature.

The trajectory of the scalar field ϕ is illustrated in Figure 1.4. At the beginning of inflation, the potential term dominates as we "slow-roll" down the potential. Neglecting the kinetic term and plugging the inflaton energy and pressure density into the second Friedmann equation (Equation 1.6), we find that in this limit the condition for inflation ($\ddot{a} > 0$) is met, and the universe expands exponentially. Toward the end of inflation, the kinetic term grows large compared to the potential and drags inflation to a halt.

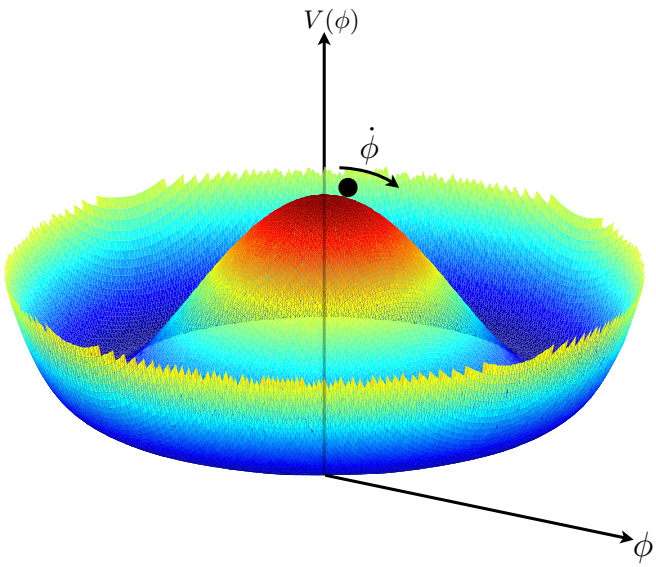


Figure 1.4: One particular inflationary model: Sombrero inflation.

This is far from the end of the story. Near the bottom of the potential, the scalar field will oscillate as it sloshes about the minimum (as will a marble in a sombrero). It is theorized that quantized oscillations of the scalar field may decay into standard particles, in a process called reheating.

Our choice in inflationary model has been arbitrary, and we emphasize that the following features are generic to slow-roll inflationary models: The first is that by requiring the potential to be nearly flat at the beginning, we can inflate the universe by many orders of magnitude. The second is that any local minimum in the potential will result in oscillations of the scalar field. The third feature is that slow-roll inflation provides a mechanism whereby perturbations can be generated across all physical scales. We explore this last point in the next section.

1.4 Perturbations in the inflationary epoch

While inflation provides a convenient picture for explaining away the horizon and flatness problems, inflation really gains traction by providing a mechanism for generating the near scale-invariant fluctuations that are observed throughout the universe. Perturbations arrise from zero-point vacuum fluctuations in the metric, which are stretched to cosmological scales by the process of inflation.

The amplitude and scale invariance of these fluctuations are critical for the growth of large-scale structure in the universe, and understanding the precise details of the origin of these fluctuations remains a critical challenge for modern cosmology. We describe in broad terms the predicted origin of perturbations from inflation. This discussion follows closely the arguments presented in Dodelson 2003.

1.4.1 Tensor perturbations

We begin by reviewing a few key principles of the quantum simple harmonic oscillator. In particular, we would like to remind ourselves how the ground state position variance is related to the frequency. We start by writing down the wave equation:

$$\ddot{x} + \omega^2 x = 0. \tag{1.36}$$

When quantized, the oscillator (famously) has non-zero zero-point energy and a non-zero ground state variance, which is given by:

$$\langle |x|^2 \rangle = \langle 0|X^\dagger X|0\rangle. \tag{1.37}$$

The position operator X can be rewritten in terms of the ladder operators:

$$X = \sqrt{\frac{1}{2\omega}}(a^\dagger + a), \tag{1.38}$$

which allows us to rewrite the variance:

$$\langle |x|^2 \rangle = \frac{1}{2\omega}\langle 0|(a + a^\dagger)(a^\dagger + a)|0\rangle. \tag{1.39}$$

As $a|0\rangle = 0$, and the ladder operators obey the commutation relation $[a, a^\dagger] = 1$, we find that:

$$\langle |x|^2 \rangle = \frac{1}{2\omega}. \tag{1.40}$$

We will use this result when deriving the variance of the tensor perturbations.

Quantum metric perturbations are the result of zero-point vacuum fluctuations, analogous to the quantized simple harmonic oscillator. The dynamics of inflation provide a mechanism whereby metric perturbations at the quantum scale are stretched to cosmological scales.

We consider a generic perturbation to the space-space component of the metric. The perturbation can be represented as a symmetric trace-free matrix (symmetric to satisfy Einstein's field equation, traceless because we are considering a perturbation). As a result, the matrix can be uniquely decomposed into scalar, vector, and tensor components. We will consider tensor perturbations first, as they are the simplest. This is because even in the presence of matter, each tensor mode k evolves independently. It can be shown that the amplitude h for some given mode k evolves according to the wave equation:

$$\ddot{h} + 2H\dot{h} + \frac{k^2}{a^2}h = 0. \tag{1.41}$$

Here derivatives are taken with respect to physical time t. This wave equation comes from considering

the space-space component of the Einstein equation, the details of which can be found in Liddle and Lyth 2000. It is illustrative to write this wave equation in terms of derivatives with respect to conformal time τ:

$$\frac{d^2h}{d\tau^2} + \frac{2}{a}\frac{da}{d\tau}\frac{dh}{d\tau} + k^2h = 0. \tag{1.42}$$

To calculate the variance of the wave equation given by 1.42 in various limiting cases, we will perform a change of variable to simplify the differential equation. Our aim is to remove the term proportional to $dh/d\tau$, allowing us to make a connection with the simple harmonic oscillator. As in Dodelson, we find that this is made possible by the following change in variable (similar in nature to changing to a co-moving quantity):

$$\tilde{h} \equiv \frac{M_{pl}}{\sqrt{2}}ah. \tag{1.43}$$

This substitution leads to the much simplified wave equation:

$$\frac{d^2\tilde{h}}{d\tau^2} + \left(k^2 - \frac{1}{a}\frac{d^2a}{d\tau^2}\right)\tilde{h} = 0. \tag{1.44}$$

Drawing from our experience with the simple harmonic oscillator, we can write down the wave equation in terms of ladder operators, just as we did in Equation 1.38 for the quantum oscillator:

$$\tilde{H} = v(k,\tau)a + v^*(k,\tau)a^\dagger, \tag{1.45}$$

with the variance given by:

$$\langle|\tilde{h}|^2\rangle = |v(k,\tau)|^2. \tag{1.46}$$

Let's consider the solution to this equation in the scenario of slow-roll inflation. In this case, $da/d\tau = a^2H \simeq -a/\tau$. We arrive at the even further simplified wave equation:

$$\frac{d^2v}{d\tau^2} + \left(k^2 - \frac{2}{\tau^2}\right)v = 0, \tag{1.47}$$

which has the solution:

$$v = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{1.48}$$

We can now find the variance for the two limiting cases described earlier. For $k\tau \gg 1$, Equation 1.48 reduces to the familiar harmonic oscillator. The variance, in analogy with Equation 1.40 is simply:

$$|v(k,\tau)|^2 = \frac{1}{2k}. \tag{1.49}$$

Similarly, for $k\tau \ll 1$:

$$|v(k,\tau)|^2 = \frac{1}{2k}\left(\frac{1}{k^2\tau^2}\right). \tag{1.50}$$

We now define the power spectrum $P_h(k)$ as:

$$P_h(k) \equiv \frac{2}{M_{pl}^2 a^2}|v(k,\tau)|^2. \tag{1.51}$$

By considering the two limiting cases above and substituting $\tau = (aH)^{-1}$, we find:

$$P_h(k) = \frac{H^2}{k^3 M_{pl}^2},\ k\tau \ll 1 \text{ (superhorizon)} \tag{1.52}$$

$$P_h(k) = \frac{1}{ka^2 M_{pl}^2},\ k\tau \gg 1 \text{ (subhorizon)}. \tag{1.53}$$

These wonderfully simple solutions have much to tell us about the evolution of quantum metric perturbations during inflation. During slow-roll inflation, $\epsilon \ll 1$ and H is nearly constant. As a result, when a wavemode exits the horizon during inflation (at quantum scales), its amplitude is frozen until the mode re-enters the horizon (at cosmological scales). In the other extreme, perturbative wavemodes much smaller than the horizon decay away as the universe expands (analogous to the cosmological redshift of the CMB). The process of inflation thus predicts a stochastic gravitational wave background at scales comparable to the cosmological horizon at the time of recombination. Without inflation, any gravitational waves present at the beginning of physical time will have long since redshifted away by recombination.

Because the tensor power spectrum is only defined up to some overall normalization relative to the scalar power spectrum, it is common to define the tensor-to-scalar ratio r, where the ratio is taken at one particular k value, typically $k = 0.002$ Mpc^{-1}. Additionally, as one might suspect from the form of Equation 1.52, departures of the power spectrum from k^{-3} proportionality are indicative of time evolution of H during inflation. This is typically parameterized as:

$$P_h(k) \propto k^{n_T - 3}. \tag{1.54}$$

The spectral index n_T is related to the slow-roll parameter ϵ by:

$$n_T = -2\epsilon. \tag{1.55}$$

Heuristically, this dependence comes about from the time evolution of the inflaton potential. If the inflaton potential is constant, then every scale is equivalent to every other scale and the spectral index is precisely zero. However, in this scenario inflation never ends. If instead the potential evolves over time, then small scales are differentiated from large scales and a spectral tilt is introduced.

As experimentalists, our goal is to measure the presence of this gravitational wave background in the early universe's history. As we will discuss in the next section, this is possible through observations of the polarization of the CMB.

1.4.2 Scalar perturbations

Much of our experience in calculating the tensor perturbation power spectra carries over when calculating equivalent expressions for scalar perturbations. Scalar perturbations are made more complicated by a coupling between the scalar field and gravity, as well as the matter content in the universe. To illustrate the general behavior of the perturbation dynamics, we will proceed assuming zero coupling and zero matter.

We begin by considering quantum fluctuations of the scalar field, given by the Klein-Gordon equation in an FRW metric:

$$\ddot{\phi} + 3H\dot{\phi} - \nabla^2\phi + V'(\phi) = 0, \tag{1.56}$$

where dots represent derivatives taken with respect to physical time and $V'(\phi) = dV/d\phi$. We can similarly write the quantum fluctuations in a first-order perturbation quantity, $\delta\phi$:

$$\ddot{\delta\phi} + 3H\dot{\delta\phi} - \nabla^2\delta\phi + V'(\delta\phi) = 0. \tag{1.57}$$

We can simplify this expression by noting that:

$$V'(\delta\phi) = V'(\phi + \delta\phi) - V'(\phi) = V''(\phi)\delta\phi. \tag{1.58}$$

Setting $m^2 \equiv V''(\phi)$ in analogy with the Klein-Gordon equation, we find:

$$\ddot{\delta\phi} + 3H\dot{\delta\phi} - \nabla^2\delta\phi + m^2\delta\phi = 0. \tag{1.59}$$

As in the case of tensor perturbations, we explore the time-evolution of one Fourier mode of wavenumber k:

$$\ddot{\delta\phi_k} + 3H\dot{\delta\phi_k} + \left(\frac{k}{a}\right)^2 \delta\phi_k + \frac{m^2}{2}\delta\phi_k = 0. \tag{1.60}$$

We can begin examining the behavior of the time evolution in the slow-roll limit. Taking into account the second slow-roll condition $V''(\phi)/V(\phi) \ll M_{pl}^2$ (Equation 1.35), we find that we can

safely neglect the last term in Equation 1.60 to find:

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \left(\frac{k}{a}\right)^2 \delta\phi_k = 0. \tag{1.61}$$

This has a very similar form to Equation 1.41. We again approximate H to be constant during inflation. The solution, up to some overall normalization, turns out to be (we refer the reader to Dodelson 2003 for details):

$$v_{\delta\phi_k}(k,\tau) = \frac{1}{k^{3/2}}(i + k\tau)e^{ik\tau}. \tag{1.62}$$

In the two limiting cases considered earlier, we find the variance to be:

$$|v_{\delta\phi}(k,\tau)|^2 = \frac{H}{k^3},\ k\tau \ll 1 \text{ (superhorizon)} \tag{1.63}$$

$$|v_{\delta\phi}(k,\tau)|^2 = \frac{H}{ka},\ k\tau \gg 1 \text{ (subhorizon)}. \tag{1.64}$$

These limiting cases reveal the same behavior as we encountered in the treatment of tensor perturbations. After perturbations exit the horizon during inflation, their time evolution stops, and their amplitude is frozen. When these modes re-enter the horizon at late times, their amplitude decays with the scale factor. This again, is in the case of zero matter and zero coupling. In reality, these scalar perturbations undergo a more complicated evolution, the result of which is observed in the acoustic peaks of the CMB.

1.5 The polarization signature of the Inflationary epoch

In this section we will describe how the presence of gravitational waves in the early universe can be inferred from mapping the polarization of the CMB. Polarization is predicted, and now measured, due to the simple combination of temperature anisotropies and Compton scattering. We will describe in detail how polarization is produced in the CMB, and how one might go about isolating the polarization anisotropy due to gravitational waves.

1.5.1 Polarization from scalar perturbations

At the time before recombination, the baryons were tightly coupled to photons. As we have discovered in the last few decades, this plasma was anisotropic at a level of roughly 10 parts per million. The mere fact that there were anisotropies in this plasma guarantees that the radiation from recombination will have some degree of polarization.

The mechanism whereby polarization is generated by scalar perturbations is illustrated in Figure

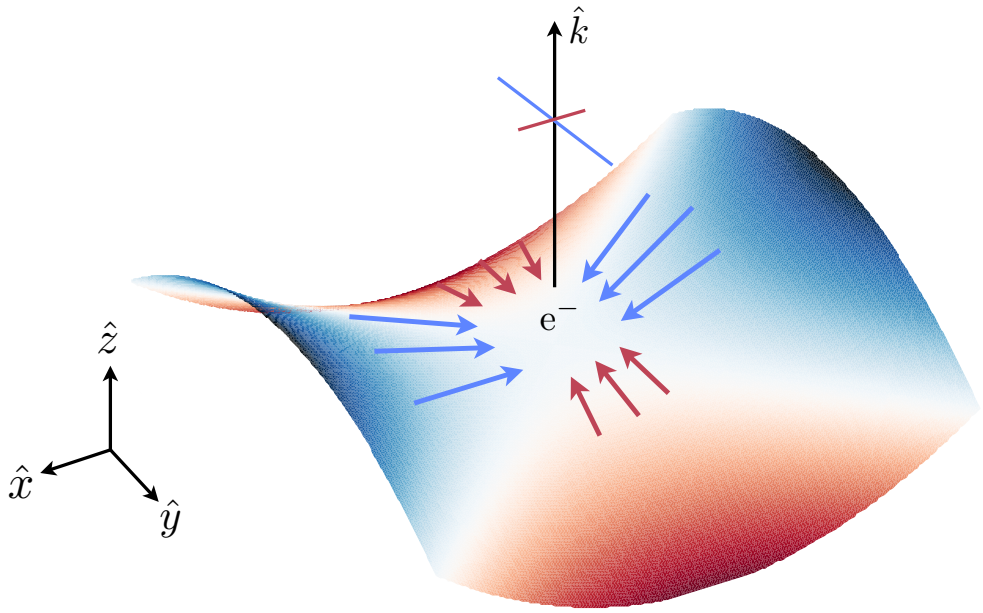


Figure 1.5: Polarization from quadrupolar anisotropy. The height of the surface in the $\hat{z}$ axis represents the amplitude of the effective potential. A free electron e^- at the center of the anisotropy will see a higher flux of photons along the $\hat{x}$ axis, resulting in a net polarization along the $\hat{y}$ axis. See text for a more complete description.

1.5. In this model, as recombination begins, electrons and protons combine, and the photon mean free path rapidly increases. We can imagine a free electron sitting in the middle of a quadrupole anisotropy. Along the $\hat{x}$ axis, the electron sits at a local minimum in the potential, and along the $\hat{y}$ axis, a local maximum. Photons emerging from this potential (represented by the blue and red arrows) will see an effective redshift or blueshift. As a result, the electron will, on average, see a higher flux along the $\hat{x}$ axis than the $\hat{y}$. This results in a net linear polarization along the $\hat{y}$ axis for a photon scattered along the $\hat{z}$ direction.

We can draw a few meaningful conclusions from this simple cartoon picture. To begin, higher and lower multipole anisotropies will not yield a net polarization. Second, the net polarization anisotropy will scale as the amplitude of the quadrupole temperature anisotropy. Additionally, we observe that polarization will only arise on scales comparable to the mean free path during recombination. At smaller scales, photons will have undergone multiple scatterings and have no net polarization. At larger scales, the number of free electrons drops considerably (due to recombination), and Compton scattering of photons stops. Finally, and perhaps most critically, we find that the net polarization orientation is tied to the direction of the scalar perturbation.

We now imagine a plane wave propagating through the early universe represented with the wave vector $\hat{k}$ (Figure 1.6). The plane wave creates higher effective potentials to the left and right of the electron (center), due to the bulk flow of photons ($\vec{v}$). The electron thus sees a quadrupolar

anisotropy, and, as we saw from Figure 1.5, this results in a net linear polarization (in this case, vertically aligned for a photon scattered out of the page). If we rotate the plane wave about $\hat{k}$, we find that the polarization state is unchanged. In this way, the polarization state is uniquely determined from the plane wave direction and amplitude; there are no additional degrees of freedom. Additionally, since the baryon-photon fluid cannot support vorticity (so $\nabla \times \vec{v} = 0$), we find that scalar perturbations give rise only to curl-free polarization fields. In analogy with electromagnetism, these are referred to as E-modes.

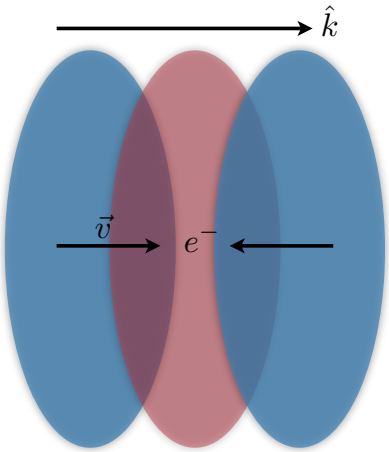


Figure 1.6: A plane wave perturbation propagating through the baryon-photon plasma. The plane wave gives rise to a bulk flow of photons represented by $\vec{v}$. The quadrupolar anisotropy seen by the electron results in a net vertical polarization of photons scattered out of the page by the electron. Under rotation about the propagation vector $\hat{k}$, the polarization state remains unchanged.

1.5.2 Polarization from tensor perturbations

The second case we will consider is polarization arising from tensor perturbations (gravitational waves). This is illustrated in Figure 1.7. A gravitational wave propagating out of the page results in oscillating compression and rarefaction in the plane of the page. We consider one slice of this propagation, corresponding to the surface of last scattering. We find the resulting polarization to be horizontal. Were we to rotate the gravitational wave vector about the axis of propagation $\hat{k}$, we would find that the polarization state rotates as well. The special symmetry exhibited by plane wave propagation is broken: it is no longer true that the polarization state is invariant under rotation of $\hat{k}$. Note that we need not rotate $\hat{k}$ in order to encounter the same problem—we only need to wait a bit longer! Were we to take a different cross-section of the same tensor perturbation, we would find a different linear polarization.

Tensor perturbations are thus responsible for producing both curl and curl-free polarization fields. Because scalar perturbations give rise to only curl-free polarization fields, performing this

component separation enables the unique identification of polarization due to tensor perturbations in the early universe. We call polarization fields that exhibit handedness B-modes. A detection of B-mode polarization in the CMB would enable us to infer the presence of a stochastic gravitational wave background present at the time of recombination. From arguments presented in Section 1.4, it follows that this would be evidence for an inflationary epoch in the early universe's history.

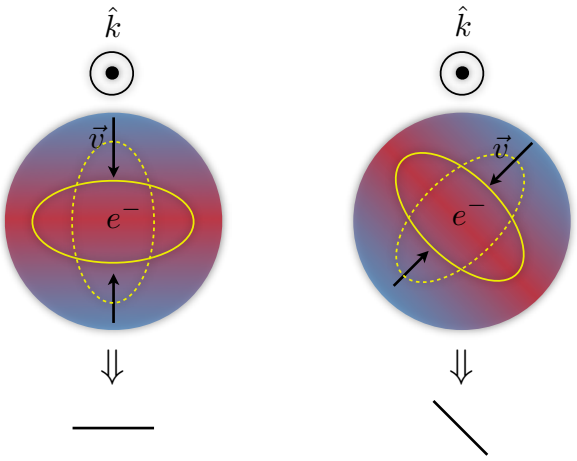


Figure 1.7: A gravitational wave perturbation propagating through the baryon-photon plasma, represented by $\hat{k}$, out of the page. The oscillatory gravitational wave h_+, to the left, gives rise to a net linear polarization, horizontally oriented (represented by the horizontal line at bottom). Unlike the case of plane wave propagation, we find that rotating about the axis of propagation rotates the resulting polarization. This is illustrated to the right. Rotating the wavevector by 45 degrees results in a corresponding rotation of the resulting polarization state. We say that polarization arising from tensor perturbations exhibits handedness.

Here we note that the B-mode polarization is the most desirable probe for tensor perturbations for the following reasons. Tensor perturbations, of course, give rise to unpolarized temperature anisotropies. So, by constraining the temperature power spectrum, it is possible to constrain the energy scale of inflation. This has been done, and to great effect. Recently, small angular scale measurements made with SPT have enabled new constraints, placing the tensor-to-scalar ratio to be $r < 0.11$ (Story et al. 2012). Ultimately, the ability to constrain r using the temperature anisotropy will be limited by degeneracy with other cosmological parameters and cosmic variance. Similarly, tensor perturbations produce E-mode polarization. E-mode polarization is ineffective at constraining r because the E-mode spectrum suffers similar degeneracies.

1.5.3 Polarization decomposition

In order to separate the polarization signature of inflation, we first need machinery to describe a polarized field as a function of position on the sky. We follow the description presented in Cabella and Kamionkowski 2004. A photon propagating along the $\hat{z}$ direction can be described by two

electric field vectors:

$$E_x = a_x \cos(\omega t - \xi_x) \tag{1.65}$$

$$E_y = a_y \cos(\omega t - \xi_y). \tag{1.66}$$

These electric field vectors can be represented by Stokes parameters as:

$$I = a_x^2 + a_y^2 \tag{1.67}$$

$$Q = a_x^2 - a_y^2 \tag{1.68}$$

$$U = 2a_x a_y \cos(\xi_x - \xi_y) \tag{1.69}$$

$$V = 2a_x a_y \sin(\xi_x - \xi_y). \tag{1.70}$$

The last term, Stokes V, reflects the circular polarization state of the radiation. As the process of Compton scattering is incapable of producing circular polarization, we ignore Stokes V. Stokes Q quantifies the polarization in the 'plus' orientation (left-right and top-bottom) while Stokes U quantifies the polarization in the 'cross' orientation (rotated 45 degrees from Q). From Q and U, we wish to calculate coordinate-independent measures of the curl- and curl-free components of the polarization field.

For our discussion, it is sufficient to proceed in a flat-sky approximation, describing the field in a local Cartesian coordinate system. This will allow us to illustrate the salient features of polarization separation before introducing spherical harmonics. As in Seljak and Zaldarriaga 1997, we find two linear combinations of Q and U that, under rotation about the origin of our local Cartesian coordinate system, transfom as:

$$(Q + iU)' = e^{2i\phi}(Q + iU) \tag{1.71}$$

$$(Q - iU)' = e^{2i\phi}(Q - iU). \tag{1.72}$$

These two complex combinations, $Q + iU$ and $Q - iU$, transform as spin-two quantities (in the two-dimensional plane) with spin-weights 2 and -2, respectively. We can construct coordinate-independent (i.e. spin-zero) quantities from these two linear combinations. Since we are dealing with spin-two quantities in two dimensions, this is possible using the spin lowering and raising operators. In the flat-sky case, these operators reduce to $S\pm = \partial_x \pm i\partial_y$ (White et al. 1999). Acting on the two spin-two quantities above, we find:

$$S_{0,-2} = (\partial_x + i\partial_y)^2(Q - iU) = Q_{xx} + 2iQ_{xy} - iU_{xx} + 2U_{xy} - Q_{yy} + iUyy \tag{1.73}$$

$$S_{0,2} = (\partial_x - i\partial_y)^2(Q + iU) = Q_{xx} - 2iQ_{xy} + iU_{xx} + 2U_{xy} - Q_{yy} - iUyy. \tag{1.74}$$

Here Q_{ij}, U_{ij} represent partial differentiation with respect to i and j. These two quantities are now coordinate independent. By rotating our coordinate system, we naturally rotate Q into U, but the above quantities are guaranteed to be invariant under those rotations. Let us test this construction by rotating by 90 degrees. As we know, the spin two quantities (Equations 1.71 and 1.72) transform simply as $(Q+iU)' = -(Q+iU)$ and $(Q-iU)' = -(Q-iU)$, respectively. Our spin zero quantities, however, will translate as:

$$S'_{0,-2} = (\partial_{x'} + i\partial_{y'})^2[(Q-iU)'] \tag{1.75}$$
$$= (\partial_{x'} + i\partial_{y'})^2[-(Q-iU)] \tag{1.76}$$
$$= -Q_{x'x'} - 2iQ_{x'y'} + iU_{x'x'} - 2U_{x'y'} + Q_{y'y'} - iUy'y' \tag{1.77}$$
$$S'_{0,2} = (\partial_{x'} - i\partial_{y'})^2[(Q+iU)'] \tag{1.78}$$
$$= (\partial_{x'} - i\partial_{y'})^2[-(Q+iU)] \tag{1.79}$$
$$= -Q_{x'x'} + 2iQ_{x'y'} - iU_{x'x'} - 2U_{x'y'} + Q_{y'y'} + iUy'y', \tag{1.80}$$

recognizing that by rotation by 90 degrees, we translate $x' = y$ and $y' = -x$. The partial derivatives thus become:

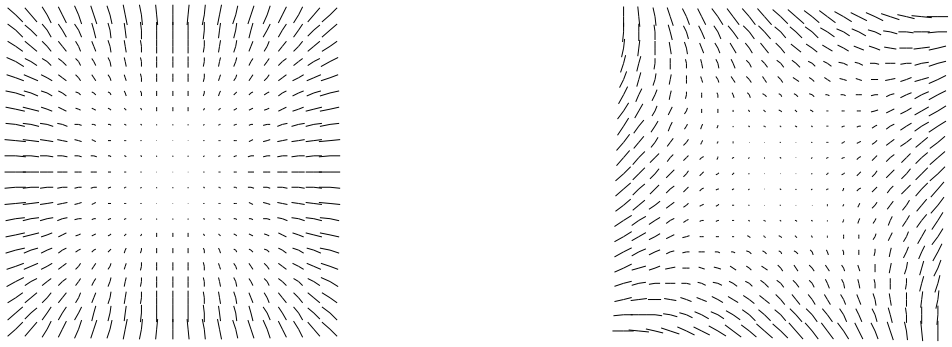
$$Q_{x'x'} = Q_{yy},\ Q_{x'y'} = -Q_{yx},\ Q_{y'y'} = Q_{xx}, \tag{1.81}$$

and similarly for U. Translating the derivatives in Equations 1.77 and 1.80 back into un-primed coordinates, we find that we recover Equations 1.73 and 1.74.

As the spin-zero quantities are rotationally invariant, so too will be any linear combination that we choose to form. Here we recall our original intent: We wish to construct parity-even and parity-odd (curl and curl-free) quantities from Q and U. We also recall that parity inversion is not the same as rotation: $S_{0,-2}$ and $S_{0,2}$ are rotationally invariant, but have ambiguous parity properties. Here we guess at a linear combination that will remain rotationally invariant and will be either parity-even or parity-odd, which we will call E and B, respectively:

$$E = -(S_{0,2} + S_{0,-2})/2 = -Q_{xx} - 2U_{xy} + Q_{yy} \tag{1.82}$$
$$B = i(S_{0,2} - S_{0,-2})/2 = U_{xx} - 2Q_{xy} - U_{yy}. \tag{1.83}$$

We test our construction by considering parity inversion about the y axis, letting $x \to -x$. Under inversion about the y axis, $U \to -U$. Acting on E and B with the parity operator Π, it is clear



(a) Pure E (b) Pure B

Figure 1.8: Polarization fields composed of pure E-modes and pure B-modes.

from Equations 1.82 and 1.83 that:

$$\Pi E = E \tag{1.84}$$

$$\Pi B = -B. \tag{1.85}$$

We arrived at the desired result: the two linear combinations of the derivatives of Q and U that we have formed obey the parity relations we require. We also note that the analog of Helmholtz decomposition holds here: any polarization field can be separated into its E and B constituents, and these components are orthogonal.

To conclude this discussion, we consider a concrete example:

$$P_1(x,y) = \alpha(x^2 - y^2) + i\beta xy \tag{1.86}$$

$$P_2(x,y) = \alpha(xy) + i\beta(x^2 - y^2). \tag{1.87}$$

These two polarization fields $P_1(x,y)$ and $P_2(x,y)$ are plotted in Figures 1.8a and 1.8b, respectively. From Figure 1.8b, we can immediately see that $P_1(x,y)$ is parity even, while $P_2(x,y)$ is parity odd. Correspondingly, if we calculate the quantities in Equations 1.82 and 1.83 we find:

$$P_{E1}(x,y) = -4\alpha - 2i\beta;\ P_{B1}(x,y) = 0; \tag{1.88}$$

$$P_{E2}(x,y) = 0;\ P_{B2}(x,y) = 4i\beta - 2\alpha. \tag{1.89}$$

As expected, the B component of P_1 is zero while the B component of P_2 is non-zero.

The derivatives corresponding to the raising and lowering operators are more easily handled in

the Fourier domain. We define the flat-sky Fourier transform of Q and U as:

$$Q(\vec{x}) = \int \tilde{Q}(\vec{u}) e^{2\pi i \vec{u}\cdot\vec{x}} d^2u \tag{1.90}$$

$$U(\vec{x}) = \int \tilde{U}(\vec{u}) e^{2\pi i \vec{u}\cdot\vec{x}} d^2u. \tag{1.91}$$

Here $\vec{x}$ is the position in the x, y plane, and $\vec{u}$ is the position in the u, v plane. We can then write our E and B constructions in terms of $\tilde{Q}$ and $\tilde{U}$. Partial derivatives with respect to x and y drop down factors of u and v respectively, resulting in:

$$E = 2\pi \int (\tilde{Q}(\vec{u})(v^2 - u^2) - 2\tilde{U}(\vec{u})(uv)) d^2u \tag{1.92}$$

$$B = 2i\pi \int (\tilde{Q}(\vec{u})(uv) - 2\tilde{U}(\vec{u})(v^2 - u^2)) d^2u. \tag{1.93}$$

Setting $|\vec{u}| = (u^2 + v^2)^{1/2}$ and $\theta = \arctan(u/v)$, we find[6]:

$$E(\vec{x}) = 2\pi \int (\tilde{Q}(\vec{u}) \cos(2\theta) + \tilde{U}(\vec{u}) \sin(2\theta)) d^2u \tag{1.94}$$

$$B(\vec{x}) = 2i\pi \int (\tilde{Q}(\vec{u}) \sin(2\theta) - \tilde{U}(\vec{u}) \cos(2\theta)) d^2u. \tag{1.95}$$

Furthermore, we can identify the Fourier transform of E and B:

$$\tilde{E}(\vec{u}) = (\tilde{Q}(\vec{u}) \cos(2\theta) + \tilde{U}(\vec{u}) \sin(2\theta)) \tag{1.96}$$

$$\tilde{B}(\vec{u}) = (\tilde{Q}(\vec{u}) \sin(2\theta) - \tilde{U}(\vec{u}) \cos(2\theta)). \tag{1.97}$$

It is thus extremely convenient to calculate E and B in the Fourier domain. In the flat-sky limit, our job here is done. We first calculate $\tilde{Q}$ and $\tilde{U}$, take linear combinations as per Equations 1.96 and 1.97, perform the inverse transforms on $\tilde{E}$ and $\tilde{B}$, and we have arrived at the desired map-space component separation.

1.5.4 Polarization on the sphere

While the BICEP2 analysis pipeline calculates E and B using the flat sky approximation, we make note of how these quantities are calculated more generally on the sphere. In the flat sky case, we constructed spin-two quantities and calculated their Fourier components. Similarly, in the spherical

[6]Note that the minus sign in front of the $\tilde{U}\cos(2\theta)$ term is a choice in convention.

case, we expand the polarization field into spin-two weighted spherical harmonics as:

$$(Q+iU)(\hat{n}) = \sum_{\ell m} a_{2,\ell m}\, {}_2Y_{\ell m}(\hat{n}) \tag{1.98}$$

$$(Q-iU)(\hat{n}) = \sum_{\ell m} a_{-2,\ell m}\, {}_{-2}Y_{\ell m}(\hat{n}). \tag{1.99}$$

Again, $\hat{n}$ is the unit vector on the sphere.

Following the same arguments as in the flat sky case, we find the following linear combinations of these expansion coefficients result in irrotational curl-free and curl components:

$$a^E_{\ell m} = -(a^{(2)}_{\ell m} + a^{(-2)}_{\ell m})/2 \tag{1.100}$$

$$a^B_{\ell m} = i(a^{(2)}_{\ell m} - a^{(-2)}_{\ell m})/2. \tag{1.101}$$

We define their angular power spectrum as:

$$\mathcal{C}^{EE}_\ell \equiv \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \left(a^{E*}_{\ell m} a^E_{\ell m}\right) \tag{1.102}$$

$$\mathcal{C}^{BB}_\ell \equiv \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \left(a^{B*}_{\ell m} a^B_{\ell m}\right). \tag{1.103}$$

These power spectra are plotted in Figure 1.9 for $r = 0.1$. We note that the peak in the primordial B-mode spectrum is predicted to occur around $\ell = 100$. This corresponds to the causal horizon scale at the time of recombination.

In large part due to theoretical development of the implications of inflation theory, we have today both a potential explanation for the origin of structure in our universe, as well as an observational window through which to explore its properties. The polarization of the CMB offers an experimental platform from which to not only test inflation theory, but to constrain the inflationary potential itself. In the next chapter, we will describe BICEP2, an instrument that has been optimized to search for the B-mode polarization of the CMB.

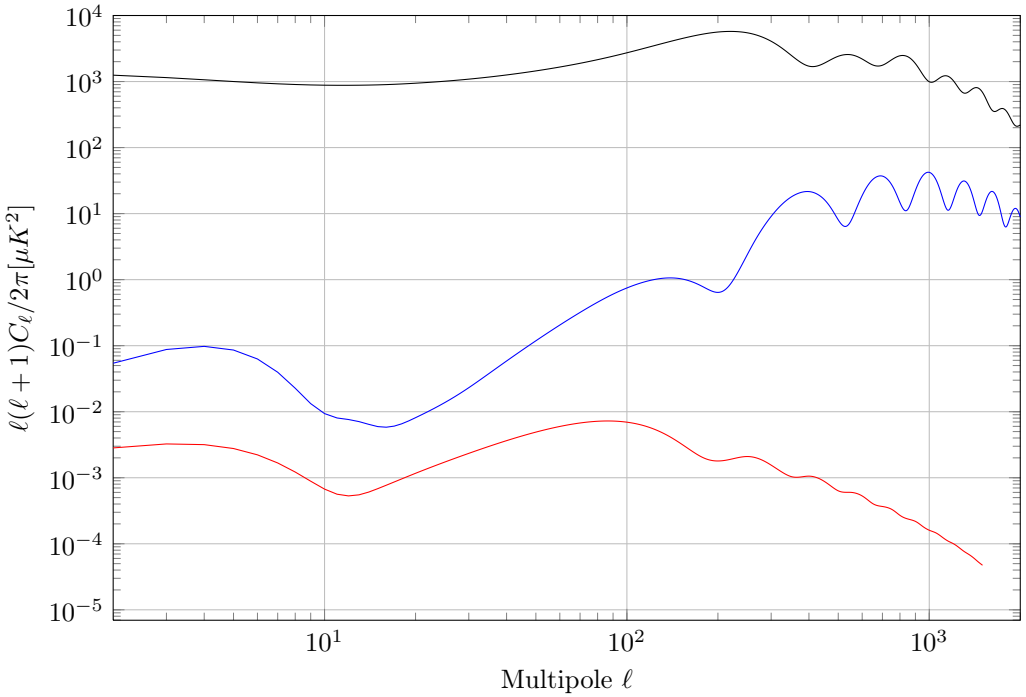


Figure 1.9: Theoretical autocorrelation power spectra for $r = 0.1$. Plotted above are the primordial signatures for $\mathcal{C}_\ell^{TT}$ (black), $\mathcal{C}_\ell^{EE}$ (blue), and $\mathcal{C}_\ell^{BB}$ (red). To date, no measurement of $\mathcal{C}_\ell^{BB}$ has been made, and its amplitude remains unknown.

Chapter 2

Instrument design

BICEP2 is the successor to the BICEP1 experiment, and one in a series of compact refractors targeting the degree-scale polarization of the CMB. Like BICEP1, BICEP2 is a purpose-built experiment for measuring B-mode polarization. At the heart of both experiments' success is cutting-edge detector technology. Indeed, what gives small, ground-based experiments like BICEP2 a competitive edge over much larger, more expensive satellite missions is their ability to rapidly field the best existing detector technology. As an experiment, BICEP2 was primarily driven by the advent of lithographed arrays of millimeter-wave polarization sensitive bolometers.

In many ways, BICEP2 builds directly on experience from BICEP1, borrowing observing strategy, design features, and even some hardware. In other ways, aspects of the BICEP1 experiment were revisited, refined, or altogether redesigned to meet the requirements of the new detector technology implemented in BICEP2.

In this chapter, we will describe the BICEP2 instrument design, calling particular attention to aspects of the experiment that greatly differ from BICEP1. We also focus our attention on design elements that could potentially introduce spurious polarization. Also, naturally, more detail and discussion is included for aspects of the experimental design in which I played a central role.

2.1 Instrument overview

What we have learned about the nature of the B-mode polarization signature in Chapter 1 directly drives the requirements of the experiment. In short, we seek to produce high-fidelity maps of the polarization of the CMB at degree scales. The angular requirement is driven by the form of the theoretical B-mode spectrum. The subhorizon and superhorizon behavior of tensor perturbations guarantees that the peak of the B-mode spectrum will occur at the causal horizon scale at the time of recombination, which today corresponds to degree-scale features on the sky. Similarly, from Equation 1.22, we conclude that an ideal instrument will achieve i) the minimum possible noise power spectrum amplitude, ii) sufficient sky coverage to suppress sample variance of the angular

scales of interest, and iii) sufficient resolution to resolve degree scales.

As we are in the business of high-fidelity polarization, we also require minimum instrumental polarization and a polarization modulation scheme. There are a wide variety of technical approaches to achieve these goals, and happily so: Various technical implementations will be susceptible to different instrument systematics. As a community, by casting a wide net in the parameter space of experimental design, we maximize the likelihood of detecting B-mode polarization.

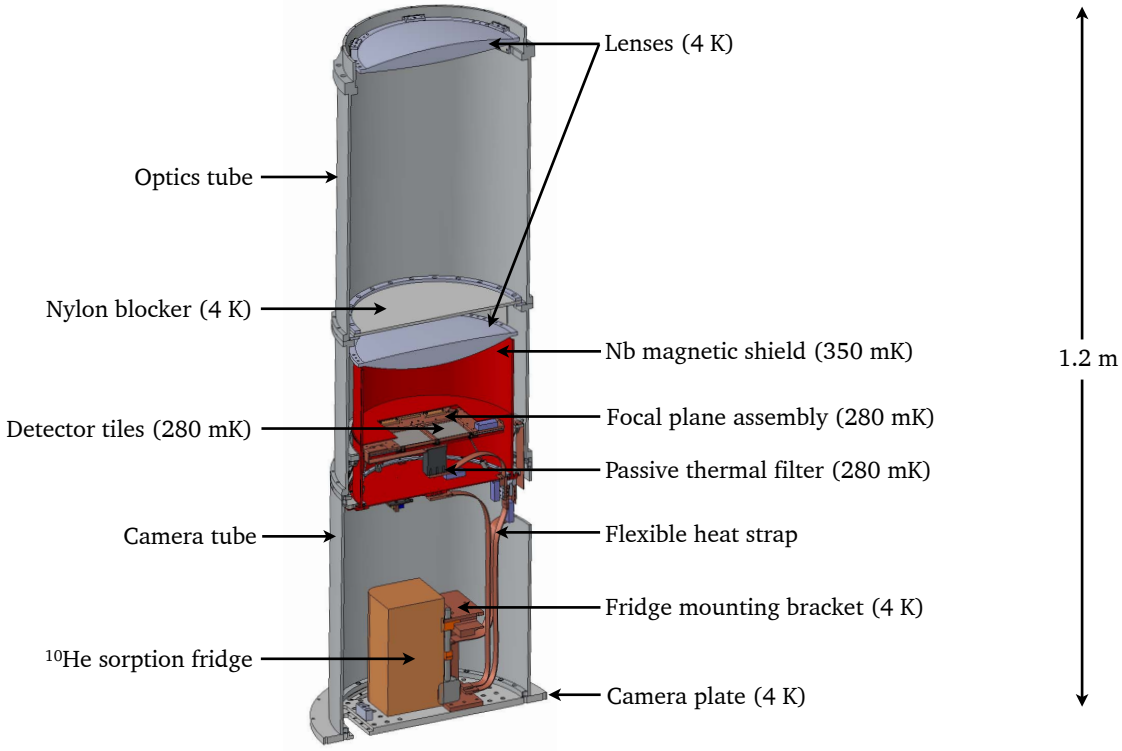


Figure 2.1: Cross-section view of the BICEP2 telescope. This cross-section includes all of the cryogenic telescope components within the vacuum vessel. Components are labeled according to the operating temperature during science observations.

The BICEP2 experimental approach is identical to that of BICEP1, but with a different technical implementation. Both experiments use the smallest aperture possible to resolve degree-scale features on the sky, while using an optical design that yields a wide field-of-view, thus preserving total light-collecting power. Additionally, we use cold, compact, on-axis refracting optics. This achieves minimal internal optical loading, low intrinsic polarization, and the ability to rotate the entire telescope about the boresight, providing polarization modulation. The detector technology is chosen to maximize the focal plane sensitivity by packing a large number of low-noise devices into a fully sampled focal plane area. Constructing a high-yield, uniform compact array comes with a large number of technical challenges, which we will discuss in subsequent sections.

Because the CMB today has a 2.7 K blackbody spectrum, reducing sources of optical loading

is key. The entire telescope is housed in a vacuum vessel, and most of the instrument operates at 4 K or below. Optical elements range in temperature from 270 K down to 4 K, with the most emissive elements at 4 K. The focal plane itself sits at 280 mK, further reducing sources of thermally, radiatively, and mechanically induced noise. The sub-Kelvin environment imposes further technical challenges that we discuss in detail below.

2.2 Optical design

The BICEP2 optical design consists of two refracting cryogenic lenses, a series of infrared (IR) blocking filters, and a vacuum window (Figure 2.2). The system is optimized for minimal instrumental polarization and minimal optical loading. Both the lens design and the filter stack were revisited for the BICEP2 experiment, in part driven by the coupling requirements to the detectors, and also in part driven by the differing frequency coverage of the experiments. (BICEP1 was a multi-frequency experiment at 100 and 150 GHz, while BICEP2 observes only at 150 GHz). We review the design of the lenses and filter stack in detail.

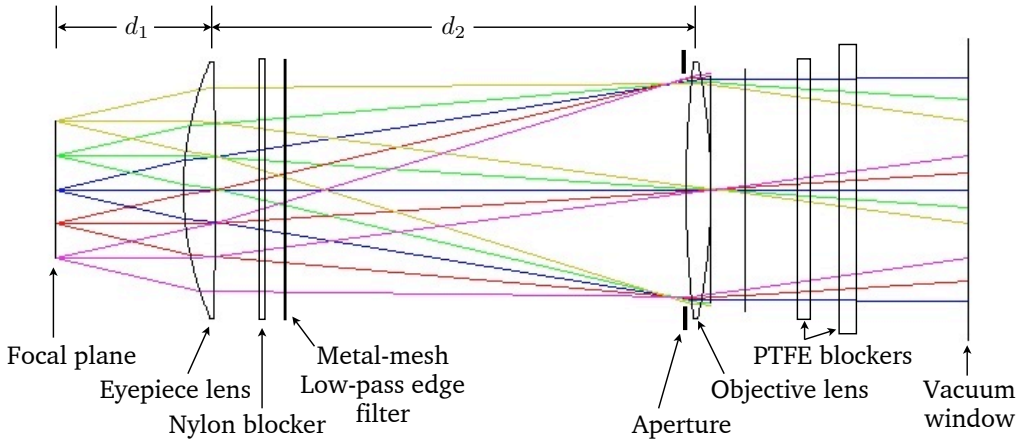


Figure 2.2: Optical layout of the BICEP2 telescope. This simple optical system consists of two refracting lenses, a series of IR filters, and a vacuum window. The eyepiece lens, objective lens, nylon and metal mesh filters, and aperture stop are all at 4.2 K. The two PTFE blockers are at roughly 40 and 100 K.

2.2.1 Lens design

The lens design was derived from BICEP1, which consisted of two refracting lenses, each roughly 30 cm in diameter. This yields an angular resolution of $1.44\lambda/D = 0.6$ degrees, corresponding roughly to the angular scales at which the B-mode spectrum is predicted to peak. Three lens systems were considered for BICEP2 to further reduce aberration in the focal plane, but the additional reflective

surfaces would introduce additional polarization-dependent reflections, and were thus disfavored. The lenses are high-density polyethylene (HDPE), cooled to the helium bath temperature, 4.2 K. HDPE is chosen for its favorable optical properties at millimeter wavelengths. The aperture stop is coincident with the first optic, making the system telecentric. The stop is also cooled to 4.2 K, providing low and stable optical loading.

The telescope $f/\#$ is chosen to match the so-called $2f\lambda$ criterionGear and Cunningham 1990, that is, chosen such that:

$$2(f/\#)\lambda = \Delta x, \tag{2.1}$$

where Δx is the pixel separation on the focal plane. For BICEP2, the pixel separation is 10 mm, and the center of the bandpass is designed for $\lambda = 2.07$ mm, yielding $f\# = 2.4$. For a 250 mm aperture, we arrive at a focal length $F = 600$ mm. Additional constraints to the focal length (and thus the plate scale) come from considerations of telecentricity and aberration. This was explored using ZEMAX[1].

Given a fixed primary aperture and focal length, there remain 10 free parameters for a simple two lens system. These are the distances between the lenses and the focal plane (d_1 and d_2 in Figure 2.2), the radii of curvature of each surface, and the conic section of each surface. We will describe the relevant constraints and optimization procedure for each of these in turn.

The distance between the focal plane and the first optic, d_1, is determined by the optical area of the focal plane, the diameter of the eyepiece lens, and the beam response of the detectors. The beam antenna response is well approximated by a Gaussian with opening half-angle of 10 degrees. The active focal plane area is roughly 180 mm in diameter. As a result, if d_1 is significantly more than 150 mm, there is significant vignetting of the beams of corner pixels by the edge of the lens. At or below 150 mm, the detectors' beams are well within the active optical area of the eyepiece lens. If d_1 is much less than 150 mm, the radius of curvature needed to maintain low aberration becomes small, making both manufacturing and anti-reflection coating problematic. Given these two competing constraints, we find an optimized distance of $d_1 \simeq 150$ mm.

In the thin-lens limit, the focal length of a telecentric optical system (and thus the plate scale) is determined only by the lens separation, d_2. Given the somewhat loose constraint of the plate scale, lens separation can be further optimized with respect to aperture illumination and aberration. In the limit of a small lens separation, aberration plays an increasingly dominant role. This is primarily because of the fact that we use a flat focal plane. Highly curved lenses are required to maintain a flat wavefront across the focal plane, and aberration quickly degrades toward the edge of the focal plane. In the limit of large lens separation, the aperture illumination (in the time reverse sense) becomes

[1] `http://www.zemax.com`

significant, both decreasing the throughput and potentially increasing far-field beam distortion, due to the fact that the system is not perfectly telecentric. Using ZEMAX simulations, we found the system to be optimized at a lens separation of 550 mm, which roughly corresponds to the aperture stop intersecting the beam at the first Airy null.

The parameters describing each lens (curvature and conic section of both sides) are optimized simultaneously. The lens surface is characterized by the surface height z above the mid-plane of the lens. The surface height, in terms of the radius of curvature r and conic section k is given by:

$$z = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}}. \tag{2.2}$$

Each lens thus has 4 degrees of freedom. These are optimized for the eyepiece and objective lenses separately in an iterative process.

We begin with a time-forward model, with incoming radiation simulated as a family of collimated rays, equally spaced between 0 and ± 8 degrees, corresponding to the full field-of-view of the telescope. The figure of merit of the optimization is the beam waist, defined as the minimum radius of a circle that encompasses all of the rays at the focal plane. In the limit that this radius is less than $\sim F\lambda/2D$, the system is said to be diffraction limited. To first order, the aberration requirement is largely determined by the curvature of the objective lens, and a relatively weak function of the details of the eyepiece lens. The curvature of the objective lens is thus largely constrained through this aberration minimization process.

After the first optimization, the objective lens details are fed into a time-reversed model, where the detectors are modeled as transmitting. The incoming field is modeled as a family of collimated ray bundles, extending from 0 to ± 10 degrees, with relative weights in the optimization according to the relative amplitude of the beam of the detectors. These ray bundles come to a focus the aperture plane, which is coincident with the front surface of the objective lens. Symmetric illumination of the aperture naturally requires that these incoming collimated ray bundles come to sharp foci in the aperture plane. This is equivalent to enforcing telecentricity, which is the requirement that rays that pass through the center of the aperture stop are parallel behind the system (in this case, behind the eyepiece lens). Using the beam waist in the aperture plane as a figure of merit, the eyepiece curvature parameters are optimized. These, in turn, are entered into the time-forward model, and the whole process is iterated.

The aberration of the system can be further reduced by floating the curvature parameters of the eyepiece lens in the time-forward model. The result is a highly non-telecentric system, resulting in asymmetric aperture illumination and far-field ellipticity. On the other hand, enforcing telecentricity at the sacrifice of aberration still results in an optical system in which a Strehl ratio of > 0.9 is achieved across the focal plane. Empirically, we find that by enforcing telecentricity, far-field beam

Parameter	Value
d_1	150 mm
d_2	550 mm
r_l, eyepiece	3643 mm
r_d, eyepiece	376.8 mm
k_l, eyepiece	532.2
k_d, eyepiece	-1.483
r_l, objective	719.3 mm
r_d, objective	1265 mm
k_l, objective	-5.727
k_d, objective	5.068

Table 2.1: Optical design parameters for BICEP2, including the lens distance d, the radius of curvature r, and the conic section k. Subscripts l and d correspond to the light and dark sides of the lenses, respectively.

distortion (primarily beam elongation) is reduced relative to the minimum aberration case. Put another way, by relaxing the requirement that the wavefront be optimally flat at the focal plane, we can make the aperture illumination more symmetric across the focal plane. By doing so, we improve the far-field beam quality. The results of this optimization process are summarized in Table 2.1.

2.2.2 Aperture stop and baffling

The aperture stop, or Lyot stop, is the optical element that defines the beam waist as it exits the telescope. The aperture stop for BICEP2 is coincident with the objective lens. This is a key advantage of the BICEP2 optical design: since the aperture stop is cooled to the liquid helium bath temperature, it provides low and stable loading. The optical system was designed so that the aperture stop would intersect the main beam close to the first null. The edge illumination varies between -12.6 and -13 dB across the focal plane (relative to the peak height). Even with this optimization, the cold stop intercepts roughly 20% of the total optical throughput.

The aperture stop itself is made from layered eccosorb[2] roughly 1.5 cm thick. The layered eccosorb was then cut to a tapered edge, and epoxied to the lens surface. The tapered edge ensures softer ringing of the sidelobes in the far-field. This is in contrast to BICEP1, where a metal knife-edge was used. Using a black absorptive aperture also ensures minimal polarization effects.

In the original BICEP2 optical design, the aperture stop was intended to sit on the light side of the objective lens. From a design perspective, this is not ideal, because the aperture stop will be radiatively coupled to warmer stages directly above the optic. As a result, we investigated the optical effects of moving the aperture stop to the dark side of the objective lens. That analysis showed that moving the aperture stop would have no discernible effect on the beam quality, so the aperture was moved to the dark side of the lens.

[2] `http:\www.eccosorb.com`

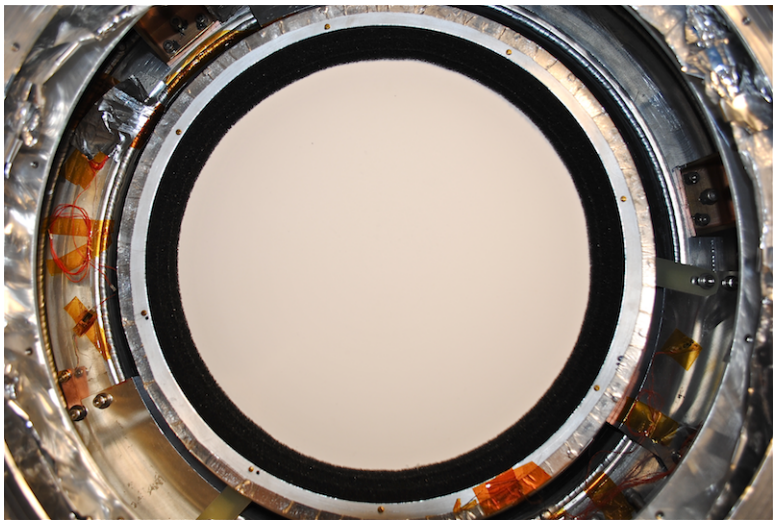


Figure 2.3: Objective lens seen from the light side in BICEP2. The aperture stop consists of a black eccosorb tapered edge. The aperture stop is on the back side of the lens shown, but is identical in construction to the baffle on the light side of the lens, seen as a black ring around the lens.

In addition to the aperture stop, there are a number of blackened surfaces at 4 K designed to terminate stray reflections from various optical surfaces. The edge of each lens is covered with an annular ring of eccosorb, similar in style to what is used for the aperture stop. The interior of the metal surfaces that support the lenses (the optics tube) are coated with carbon-loaded Stycast (colloquially known as Bock black). The textured blackened surface was measured to be $> 95\%$ absorptive.

2.2.3 Filter stack

The filter stack in BICEP2 consists of a Zotefoam vacuum window, two Teflon (PTFE) blockers at 40 and 100 K, a nylon blocker at 4 K, and a metal mesh filter produced by Peter Ade. All are anti-reflection coated for 150 GHz. The vacuum window is constructed from Zotefoam PPA30 extruded polypropylene foam, of the same variety used in BICEP1. The two Teflon blockers are used primarily for rejecting IR radiation that would otherwise compromise both detector noise performance as well as cryogenic performance. Teflon is largely opaque in the IR, but minimally emissive at millimeter wavelengths. Assuming 2% emissivity in-band, the Teflon blockers at 40 and 100 K are predicted to give 0.2 and 0.4 pW optical loading in-band, respectively. Because of the poor thermal conductivity of Teflon, the center of the filter may be substantially warmer. Even in this case however, the in-band emissivity is well within the requirements of the experiment.

To further reduce loading, a 6 mm thick nylon filter is used directly above the eyepiece lens. Nylon substantially cuts down on the far-IR loading on sub-Kelvin stages, but at the sacrifice of reduced optical throughput and added optical load. Since the filter is at 4 K, the calculated optical

load is much less than 0.1 pW, even if an unreasonably high in-band emissivity is assumed. There is some reduced optical efficiency due to the added reflective surfaces and in-band absorption, but this has been measured to be less than 2% at 150 GHz.

The final filter in the optical chain is a metal mesh band-defining filter with a tuned frequency cutoff of 8.3 icm. The use of this filter was motivated by an observed percent-level high-frequency spectral leak in the detectors. This measured "blue leak" was partially mitigated by changes to the detector design before the BICEP2 science-grade detectors were fabricated. With the combination of these detector design changes and the presence of the metal mesh filter, this out-of-band coupling has been largely mitigated.

2.2.4 Anti-reflection coating

All of the optical elements (except for the vacuum window) were anti-reflection coated using porous Teflon, manufactured under the product name Mupor[3]. The porous Teflon can be tuned in both thickness and density to match the $\lambda/4$ criterion for impedance matching.

The anti-reflection coating was heat-bonded using a vacuum bagging procedure. Each optical element was placed in an inverted mold and coated with a thin low-density polyethylene (LDPE) layer and the porous teflon anti-reflection coating. The mold, together with the optical element and the anti-reflection coating, was enclosed in a silicone bag, which was then placed under vacuum and inserted into a convection oven (Figure 2.4). The vacuum bag ensured isobaric pressure was applied evenly to the surface during the heat bonding.

During the first few attempts at anti-reflection coating in this way, it was discovered that internal stresses significantly deformed the optical elements under thermal cycling, particularly for the HDPE lenses. As a result, a great deal of effort was made to anneal the optical elements during the manufacturing process. After following a two-stage annealing schedule, we found the lens shape to be highly repeatable and unchanged by the anti-reflection coating procedure.

2.3 Detector design

The beating heart of BICEP2 is an array of 512 Transition-Edge Sensor (TES) bolometers, forming what was, at the time of deployment, one of the most sensitive microwave polarimeters ever built. TES technology has allowed not only improved per-detector sensitivity, but also the ability to multiplex many detectors simultaneously, making possible focal planes with hundreds, rather than dozens, of detectors.

In this section, we will provide a broad overview of the operating principles of the BICEP2 focal plane, including a review of our TES bolometers, the phased-array antennas, and the lithography

[3] `http://www.porex.com`

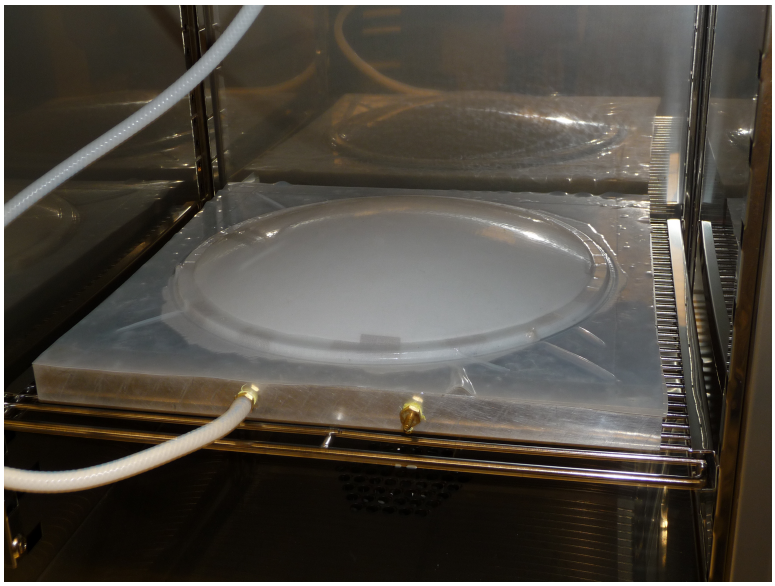


Figure 2.4: Anti-reflection coating of the BICEP2 lenses. The lens is placed in an inverted mold, coated with LDPE and the anti-reflection film, and placed in a vacuum bag. The entire assembly is then pulled under vacuum and placed in a convection oven for heat bonding.

process used to fabricate the devices.

2.3.1 Transition-Edge Sensors bolometers

The principle of operation for any bolometer is simple: An absorptive element is coupled to a thermally isolated temperature-sensitive resistor. Incident radiation heats the absorber, which in turn heats the resistor, which is then measured as a change in current. The TES bolometer is no different. The temperature-sensitive resistor is a superconducting film, biased onto its superconducting transition. Small changes in temperature induce corresponding changes in resistance. The previous generation of bolometers used in BICEP1 (called NTD bolometers) used semiconducting films to achieve the same function.

The bolometer must be minimally thermally coupled to the rest of the system, but must also be able to dissipate power back to the bath. For this reason, the bolometer is formed into an "island", which is thermally coupled to the bath temperature, $T_{\rm bath}$ through a weak heat link of conductivity G (Figure 2.6). The device thus also naturally has some thermal time constant, determined by the conductivity G and the heat capacity of the island, C.

The TES bolometer offers several key advantages over its semi-conducting predecessor. First, when optimally biased, the resistance of the TES is a steep function of temperature, the result of which is very high responsivity. The second advantage is that TES devices can be read out using SQUID amplifiers rather than JFETs, which is a critical consideration when one considers multiplexing a large number of these devices on a single focal plane. A third advantage is that the

thermal time constant of the TES devices is much faster than for NTDs due to the much smaller heat capacity of the superconducting device. NTD-based detectors often require complicated modeling to fully characterize the temporal transfer function of the device. There is no such complication with TESs: The devices are sufficiently fast that the temporal transfer function can often be entirely ignored in analysis[4]. Finally, arrays of TESs can be fully lithographed onto silicon wafers, avoiding the hand assembly required for feedhorn-coupled NTDs.

The TES can be stably biased onto its superconducting transition using negative electro-thermal feedback. We consider a device with a thin film of superconducting metal with a transition temperature T_c, typically of order ~ 500 mK. A voltage bias is applied to keep the film between the superconducting and normal states. An increase in resistance results in a corresponding increase in Joule power, which in turn, drives up the temperature. As the temperature increases, the bias current decreases, along with the Joule heating, resulting in a negative feedback loop. Expressed more formally, electro-thermal feedback is the result of two coupled differential equations. Ignoring noise, we can write the thermal differential equation:

$$C\frac{dT}{dt} = -P_G + P_{\rm J} + P. \tag{2.3}$$

Here C is the heat capacity, P_G is the power lost to the bath, $P_{\rm J}$ is the Joule power, and P is the incident optical power. Correspondingly, the characteristic electrical differential equation is:

$$L\frac{dI}{dt} = V_{\rm bias} - IR_{\rm sh} - IR(T,I), \tag{2.4}$$

where L is the electrical inductance, $V_{\rm bias}$ is the bias voltage, I is the current, $R_{\rm sh}$ is the shunt resistance (which may also contain some parasitic resistance), and $R(I,T)$ is the temperature-dependent resistance.

The coupling between these two differential equations comes about from the Joule power term, which is:

$$P_{\rm J} = V_{\rm bias}^2/R(T,I). \tag{2.5}$$

It is useful to examine the steady-state behavior of these expressions. Setting the time derivative terms to zero, and solving for the optical power P:

$$P = P_G - \frac{V_{\rm bias}^2 I}{V_{\rm bias} - IR_{\rm sh}} \tag{2.6}$$

We conclude from this expression that in the steady-state case, if we know the various parameters of

[4]The speed of the TES can also be a nuisance. TES bolometers can become unstable when the thermal timeconstant and electrical timeconstant overlap. There is no such concern with NTDs.

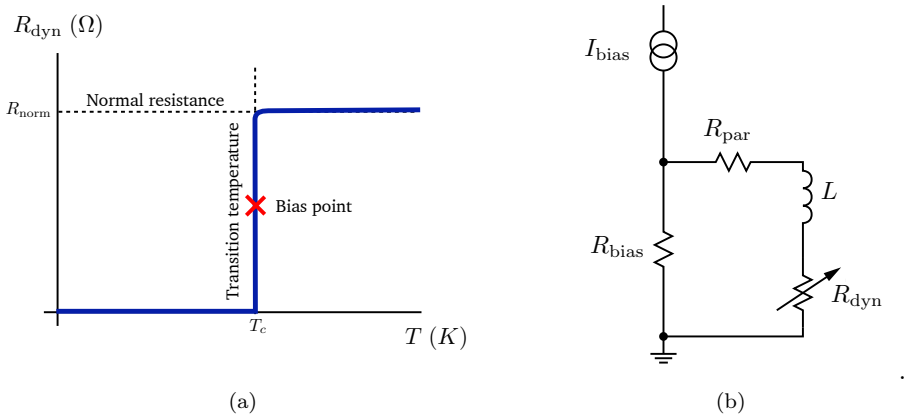


Figure 2.5: *Left:* Resistance versus temperature model for an idealized TES detector. During operation, a voltage bias, $V_{\rm bias} = I_{\rm bias} R_{\rm bias}$, is applied to lock the detector onto its superconducting transition. Small changes in temperature due to incoming radiation result in dramatic changes in resistance. *Right:* A simple TES bias circuit, adapted from Irwin and Hilton 2005. A Thevenin equivalent circuit can be drawn, substituting $R_{\rm sh} = R_{\rm bias} + R_{\rm par}$.

our system $(G, C, R_{\rm sh}, T_{\rm base}, V_{\rm bias})$, we can directly infer the optical power incident on the device by reading the applied current, I. In the dynamic case, electrothermal feedback is used to 'lock' onto the superconducting transition of the TES, balancing the left and right-hand-sides of this expression.

The TES device readout takes advantage of this electro-thermal feedback mechanism. A change in optical power on the TES is compensated by a change in current (due to the voltage bias and the changing resistance) to maintain a steady lock point on the superconducting transition. The applied feedback current is read out and digitized. Using a series of calibration steps, the change in current can be referenced back to the equivalent change in optical power, and a corresponding brightness fluctuation on the sky.

The transfer function of the device is primarily determined by the thermal and electrical time-constants. As mentioned previously, the bolometer is suspended on a thermally isolated island, connected to the heat bath via thin, thermally conductive legs. The characteristic thermal time constant is simply:

$$\tau = \frac{C}{G}. \tag{2.7}$$

In the case of BICEP2, the islands are etched from a substrate, leaving thin legs to suspend the island. Both the C and G are tunable by adjusting the geometry of the device, which is necessary to optimize the noise performance. The electrical time constant is given by:

$$\tau_{\rm el} = \frac{L}{R_{\rm sh} + R_{\rm dyn}}, \tag{2.8}$$

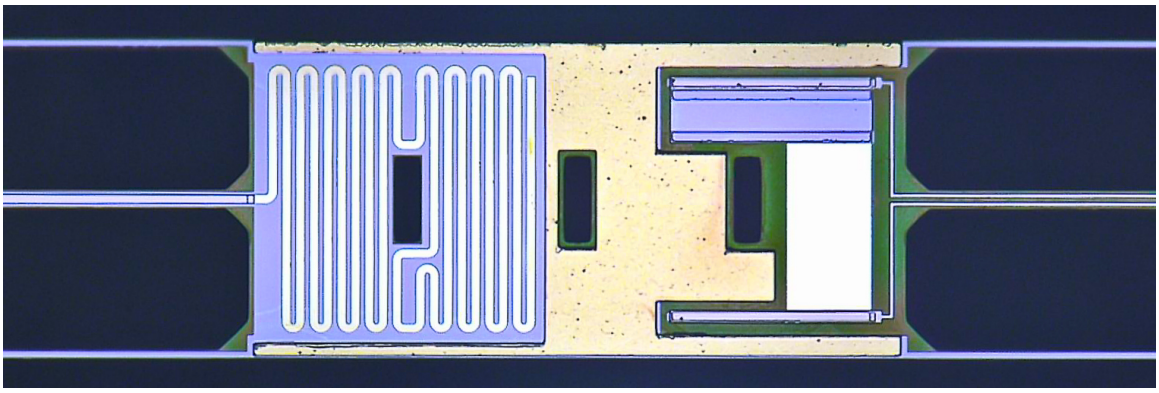


Figure 2.6: Microscope photograph of the detector island of one BICEP2 TES bolometer. The island is supported by six silicon nitride legs (left and right). Microwave radiation enters on the left-hand side of the figure and is terminated on a gold meander. A voltage is applied to the device via the micro-strip lines entering on the right. The TES consists of an aluminum film and titanium film in series, seen as the blue (Ti) and white rectangles (Al) on the right-hand side of the island. The three dark rectangles in the center of the island are through holes used to improve the island release process.

where $R_{\rm dyn}$ is the dynamic resistance $\partial V/\partial I$ at fixed temperature T_0. The typical dynamic resistance for a BICEP2 TES is ~ 0.1 mΩ.

The dominant sources of noise of a TES bolometer are photon noise, Johnson noise, thermal noise, and the so-called "excess" noise, the last of which is today not fully understood. A complete treatment of TES noise components can be found in Irwin and Hilton 2005. A complete description of the measured TES noise components of the BICEP2 devices can be found in Brevik 2012 and references therein. Given the wealth of resources that comprehensively cover TES noise, here we provide only a heuristic description of these various noise components.

Photon noise, as its name implies, results from noise induced by the optical load due to the random arrival rate of photons. In the limit of high occupation number (for bright, thermalized sources) the noise is proportional to the source temperature. Johnson noise is current noise arising from thermal motion of electrons within any resistive device. It is spectrally white, and is rolled off at high frequencies by the electrical time constant. At low frequencies, Johnson noise suppressed by electrothermal feedback. In this context, thermal noise refers to the thermal analogue of Johnson noise, corresponding to shot noise on the phonons that carry power between the bolometer island and the thermal bath. Correspondingly, it is rolled off at high frequencies by the electro-thermal feedback time constant of the device. The final significant noise component is TES excess electrical noise, so-called because it is not predicted by present theory. While not presently understood, it does exhibit dependence on both the lock point on the TES transition as well as on the island geometry.

TES devices become non-linear, and indeed saturate, when the optical power P dominates over the Joule power P_J from the bias current. At some point, the optical power is sufficient to drive the

device normal. The saturation power of a TES is given by:

$$P_{\rm sat} = P_G(T) - \left(\frac{V}{R_{\rm sh} + R_N}\right)^2 R_N, \tag{2.9}$$

where R_N is the normal resistance, V is the voltage across the bolometer, and $P_G(T)$ is the power flowing to the bath. Higher conductivity thermal links of the island to the bath can raise the saturation power, but at the cost of sensitivity. For BICEP2, it is critical to be able to achieve a dynamic range spanning many orders of magnitude to be able to observe our science target (at 3 K), as well as ground-based microwave calibrators (which can be $\sim 1 \times 10^6$ K). As a result, the TESs in BICEP2 consist of two superconductors in series: the first is made of titanium and is used for science observations, with a transition temperature of roughly 500 mK. The second device in series is an aluminum film with a superconducting temperature of around $1.3-1.4$ K. This is used for calibration measurements, using both microwave sources and bright astronomical calibration sources.

2.3.2 Phased-array antennas

Radiation is coupled to the TES devices using phased-array antennas, which are coupled to the TES device via a micro-strip summing tree. The BICEP2 antenna consists of an array of "cells" of slot dipoles, arranged in a 12×12 grid, which are coherently summed through a binary[5] summing tree. Each cell, or sub-radiator, consists of 4 slots, as illustrated in Figure 2.7. The slots, along with the micro-strip summing tree, are lithographically patterned onto a substrate. The micro-strip is designed such that the optical path length between the detector and any given slot is the same. The microwave power is received on the TES island by a gold meander.

The BICEP2 detectors measure two linear polarizations simultaneously. Two arrays of slot antennas, together with two interleaved summing trees, couple linear polarizations A and B to two independent TES bolometers. The beam-forming arrays are designed to be perfectly coincident on the sky.

Previous generations of feedhorn-coupled bolometers relied on optical filters, as well as the waveguide cutoff, to define the spectral band. Since BICEP2 is antenna-coupled, this requires the spectral passband to be defined on-chip. The BICEP2 passband is defined by an inline lumped element filter, lithographed onto the substrate together with the rest of the detector. The filter is a 3rd-order Chebyshev filter, formed from stub capacitors and coplanar waveguide inductors. Microwave power passes through this filter at the "trunk" of the summing tree before the power is terminated onto the TES bolometer.

[5]The BICEP2 summing tree is not precisely binary, as it has 12 rather than 16 divisions.

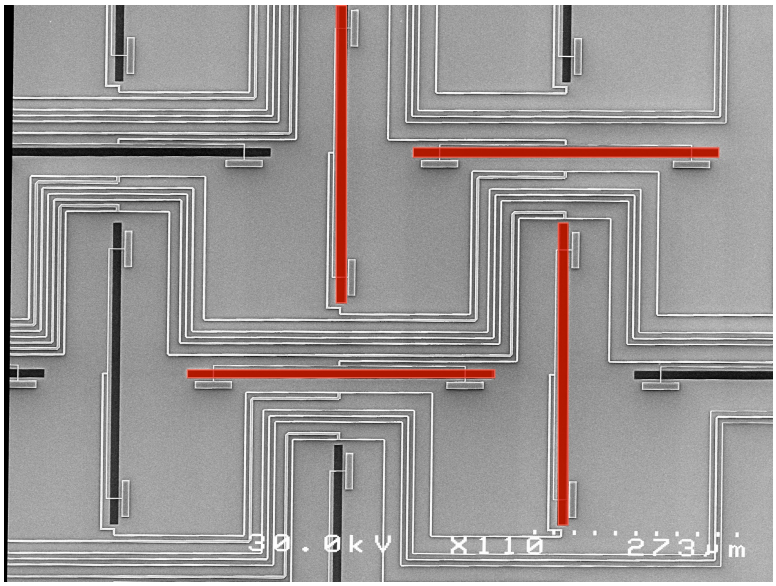


Figure 2.7: Slot array and summing tree of the BICEP2 devices. The image of this detector, created using scanning electron microscope data (SEM), shows a small area of the total 12×12 array of sub-radiators. The sub-radiator, consisting of four slots (two horizontal and two vertical), is highlighted in red. The slot taps are formed by the white rectangles and the ground plane, which are coherently summed via a micro-strip summing tree (thin white lines) to create a beam-forming array.

2.3.3 Lithography process

The BICEP2 detector tiles were fabricated by the Micro-Devices Laboratory (MDL), a division of the Jet Propulsion Laboratory (JPL) by Anthony Turner and Joseph Bonetti. The process is described in detail in Bonetti et al. 2009; Bonetti et al. 2009; Bonetti et al. 2012. We give a general overview of the process used for BICEP2 here.

The process begins with a 4 inch diameter silicon wafer. A 1 μm silicon nitride (Si_3N_4) top layer is grown via a low-pressure chemical vaporization deposition. This top layer serves as the material for the thermally isolating legs as well as the TES island itself.

The Al TES is deposited first, using electron beam evaporation and patterned by liftoff. The Al TES is protected from later stages of lithography with a thin layer of sputtered silicon dioxide (SiO_2). The protective SiO_2 is selectively etched to allow electrical contact of the Al TES with the rest of the circuit. The Ti TES is deposited in the second step using a 6 inch diameter sputtering gun and patterned with inductively coupled plasma (ICP) etching.

After the TES films have been deposited, the niobium (Nb) ground plane is patterned, forming the slot dipoles as well as the ground plane shield. Historically, this step in the fabrication process has been somewhat variable: Devices have been tested where the ground plane has been patterned using both etching and liftoff processes. For BICEP2, liftoff was chosen, as it empirically yielded higher uniformity in device parameters. There is some evidence, however, that the liftoff process

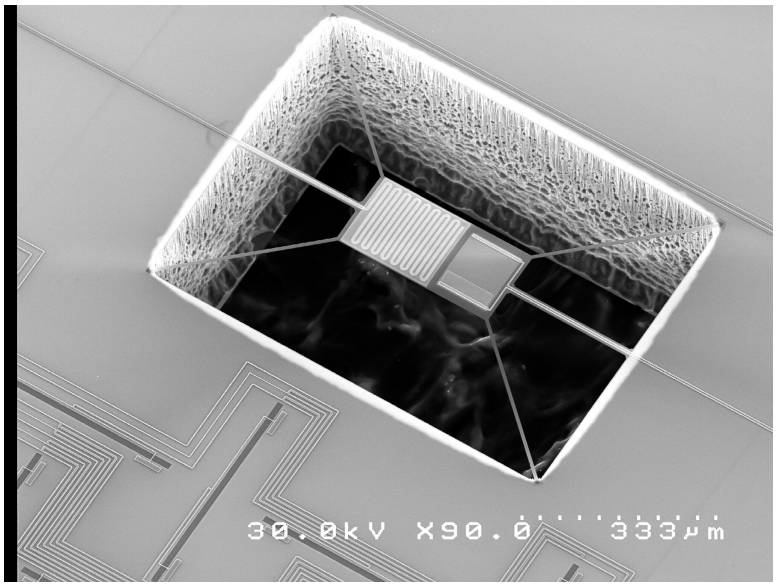


Figure 2.8: Lithography of an early generation BICEP2 device. The TES island is supported on thin Si_3N_4 legs. The device is released from the substrate via a deep reactive ion etcher and undercut using XeF_2 gas.

may adversely affect the beam quality.

A layer of SiO_2 serves as the interlayer dielectric (ILD) between the ground plane and the wiring layer deposited in later steps. The ILD is deposited in two steps via RF sputtering. The two-step deposition, with an intermediate polishing and cleaning step, helps reduce the number of pinhole shorts in the dielectric and improves the thickness uniformity. Various microwave properties, including the spectral bandpass and microwave phase velocity, are sensitive to the ILD thickness. As a result, non-uniformity can directly contribute to spectral and optical mismatch between detectors within a polarized pair. With the ILD deposited, vias are etched to make electrical contact between the TESs and the Nb wiring (top) layer.

The gold meander that acts as the terminating resistor for the antenna is deposited next using electron-beam evaporation and liftoff. Additionally, 2.5 μm of gold is deposited in un-occupied real estate on the TES island. The purpose of this additional gold is to increase the heat capacity of the island, thereby avoiding electro-thermal feedback oscillation and attenuating noise outside of the readout electronics bandwidth.

The Nb top layer is then deposited, using the same processing steps as for the Nb ground plane. Here again, there has been some variability in the process. For BICEP2, the top layer was patterned in a liftoff process. The Nb top layer includes the wiring layer, including the slot taps, the summing tree, and the lumped-element filter. Additionally, the wiring bus for the TES readout is patterned in this step.

The final step in the lithography process is to "release" the TES island. The surface features of

the SiO_2 and Si_3N_4 legs and island are patterned with an ICP etch. Next, the substrate itself is etched, leaving the TES island thermally isolated. The substrate is treated with a "deep trench" etch, which bores a deep, smooth-walled hole in the silicon using a deep reactive ion etcher. The island is then undercut using XeF_2 gas. Since the undercutting process determines the island volume, it also determines the device heat capacity. For this reason, there are several additional holes in the center of the island that assist in achieving uniformity in the undercutting process. The end product of this process is pictured in Figure 2.8 for an early-generation BICEP2 device imaged with an SEM.

2.4 Focal plane and readout electronics

A key advantage of antenna-coupled TES detectors is that arrays of these detectors can be fabricated completely using photolithography. The BICEP2 focal plane consists of four silicon wafers or "tiles," each with 8×8 spatial pixels, each with two polarization sensitive detectors, for a total of 512 devices. The large detector count is made possible using a multiplexed readout and a novel focal plane architecture. We first describe the readout scheme for the BICEP2 detectors, then discuss the implementation and focal plane mechanical structure.

2.4.1 SQUID readout

Changes in applied current to maintain the TES lock point are read out using superconducting quantum interference devices (SQUIDs), which are superconducting magnetometers. They are turned into ammeters by coupling to the SQUID the magnetic flux from an inductor through which the current to be sensed flows. At this point, we take a small detour to detail the principle of operation of a SQUID.

The SQUID operates by combining the quantum mechanical effects of flux quantization and Josephson tunneling. The DC SQUID consists of a superconducting loop, interrupted by asymmetrical Josephson junctions in each of the two loop branches. The superconducting loop is penetrated by some integer number of flux quanta ($n\Phi_0$). When a magnetic field is applied, a screening current develops to maintain an integer number of flux quanta. A bias current is applied to the SQUID so that the critical current Josephson junction is exceeded and a voltage develops across the SQUID. When this bias current is applied, a screening current can be measured as a change in voltage across the device. In this way, the SQUID can be used to measure the applied magnetic field.

In practice, the magnetic field is applied to the SQUID via an inductively-coupled "input coil." When a current is applied to the input coil, a magnetic field is applied to the SQUID resulting in a change in the voltage across the SQUID. Since the screening current drops to zero when the applied magnetic field is equal to an integer number of flux quanta, the SQUID voltage is a periodic function of the applied input coil current, with periodicity $I_0 = \Phi_0/M$, where M is the mutual inductance

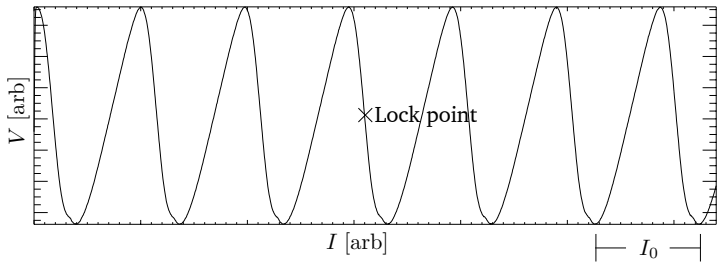


Figure 2.9: Example SQUID $I-V$ curve. The voltage is a periodic function of the applied current with periodicity I_0. The response of the device can be linearized by locking onto a steep, smooth portion of the $I-V$ curve. This particular $I-V$ curve is taken from one of BICEP2's second stage amplifiers.

between the input coil and the SQUID. A typical BICEP2 $I-V$ curve is plotted in Figure 2.9.

Like the TES, the gain of the SQUID can be linearized by locking onto the device where the current responsivity is large. In this regime, small changes in the applied field result in large changes in the measured voltage. The feedback is established through a "flux compensation" (FC) loop. Each SQUID is coupled to two inductors. The first is the input coil, tied to the TES. Changes in TES current induce changes in the applied field to the SQUID. The second is the feedback coil, which is tied to a flux compensation loop that increases or decreases the applied field necessary to maintain the same voltage across the SQUID. The response is thus nulled, and the signal is measured as the current necessary to maintain a constant voltage across the SQUID. The SQUID readout for BICEP2 is fabricated by NIST, led by Kent Irwin. Details of the fabrication and characterization of the devices can be found in Chervenak et al. 1999, de Korte et al. 2003, and references therein.

The gain (or change in applied current at the SQUID to the changing current at the TES) is not large enough to read out the SQUIDs directly with room-temperature electronics. As a result, BICEP2 employs three amplification stages with SQUIDs. The amplifier chain is inherently tied to the multiplexing architecture, which we will describe in Section 2.4.2.

The first and second SQUID amplifier stages consist of individual SQUIDs, while the third stage is an array of 100 SQUIDs connected in series, as an aptly-named SQUID series array (SSA). Current is applied to the SSA through many successive inductive coils that couple to each SQUID in series. The voltage is then measured across all of the SQUIDs in the array. This yields a highly amplified signal, which can then be read out using room-temperature electronics.

As described in de Korte et al. 2003, the SQUID readout system is very high bandwidth. This allows signal multiplexing for large format arrays, as we describe in the next section.

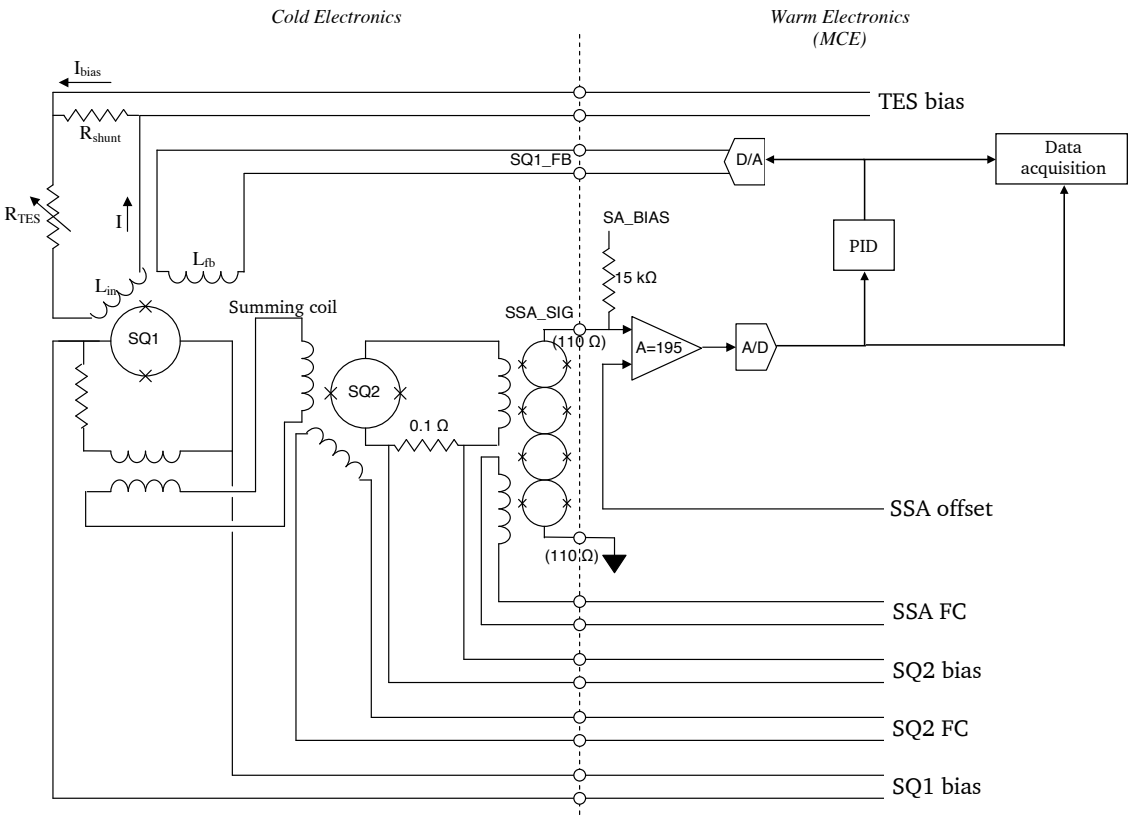


Figure 2.10: Schematic of the time-domain multiplexing readout system. The left-hand side of the figure illustrates the cryogenic electronic components, while the right-hand illustrates the room-temperature components. Acronyms used include SQUID series array (SSA), first-stage SQUID (SQ1), second-stage SQUID (SQ2), analog-to-digital (A/D), digital-to-analog (D/A), and flux compensation (FC). Multiplexing is achieved by switching the SQ1 bias lines at 25 kHz. Figure adapted from `http://www.phas.ubc.ca/~mce/`.

2.4.2 Time-domain multiplexer

The BICEP2 time-domain multiplexing (TDM) scheme operates by reading out many detectors on the same circuit successively in time. Each TES detector is coupled to a first-stage SQUID (SQ1) by an inductor, termed the "input coil." Groups of 33 SQ1s (columns) are then read out by a single second-stage SQUID (SQ2). This is made possible by applying a square wave bias to the SQ1 bias line. Each of the SQ1s in a column is turned on in succession, and the signal from each of the SQ1s is thus separated in time. The SQ1 bias lines are called row select lines, and are shared between all columns in the focal plane. The SQ1 signals are routed to the SQ2 via a large summing coil, which inductively couples the current flowing through each of the SQ1s within the column to its shared SQ2. There are a total of 16 columns on the focal plane.

After the signal propagates to the SQ2, it travels down superconducting NbTi cables to the SSAs at 4 K. Here it is further amplified, and then carried via Manganin wires to the room temperature electronics. To disentangle the multiplexed signal, one need only assign a column (corresponding

to a particular cable pair) and row (corresponding to a particular division in time) to reconstruct the signal of each detector on the focal plane. The operational multiplexing frequency is 25 kHz, meaning that every detector is visited at that rate.

The room temperature electronic readout is provided by the Multi-Channel Electronics (MCE) system provided by the University of British Columbia (Battistelli et al. 2008). This system has been used by a large number of experiments, including SCUBA2, Keck, and ACT. A basic schematic of the MCE can be seen in Figure 2.10, and a full functional description can be found at `http://www.phas.ubc.ca/~mce/`.

2.4.3 Focal plane architecture

The BICEP2 focal plane package consists of 4 detector tiles, 16 SQUID multiplexing (or MUX) chips, 16 Nyquist chips, and the wiring bus to carry the signal traces from the detector tiles through SQ1 and SQ2. The focal plane also has a number of mechanical features to facilitate the optical and thermal performance of the devices.

The detector tiles are mounted, together with anti-reflection tiles, onto a thick piece of gold-plated copper (called the detector plate) as illustrated in Figure 2.11. Light enters through 4 square holes in the detector plate. The anti-reflection tiles are tuned for impedance matching of the silicon substrate to free space. Both are held to the detector plate using spring-loaded clips. The tiles are pin aligned, and are allowed to slide under the clips during thermal cycling. The multi-layer printed circuit board (PCB) that routes the TES signals to the MUX chips is also mounted onto the detector plate, occupying the outer radius of the focal plane. The detector tiles are wire-bonded to the PCB using Al wire bonds. The top layer of the PCB is Al plated, and, as a result, the signal bus between the TES and the SQ1 is largely superconducting.

The MUX chips and Nyquist chips are together mounted on a ceramic carrier, which is in turn epoxied to the PCB and wire bonded. The signal traces are then routed to six 37-way connectors. The multiplexing scheme shares the row select lines between MUX columns. This requires a signal trace "race track", which connects each of the columns in series. To reduce the loop area, the race track uses vertically overlapping send and return traces.

Mounted to the dark side of the detector plate is a contiguous plate of niobium. This plate serves as both an optical backshort as well as a magnetic shield. As an optical backshort, the plate is placed at a distance of $\lambda/2$ from the detector tile. This boundary condition improves the optical efficiency of the device. With a transition temperature of 9 K, the superconducting plate is extremely effective at shielding both the TESs and the SQUIDs via the Meissner effect. Also, with a transition temperature above the helium bath temperature, the magnetic shield remains superconducting through cryogenic fridge cycles. The magnetic shielding design is described in further detail in Section 2.6.

A focus of the design work of the BICEP2 focal plane was to have single-point cooling for the

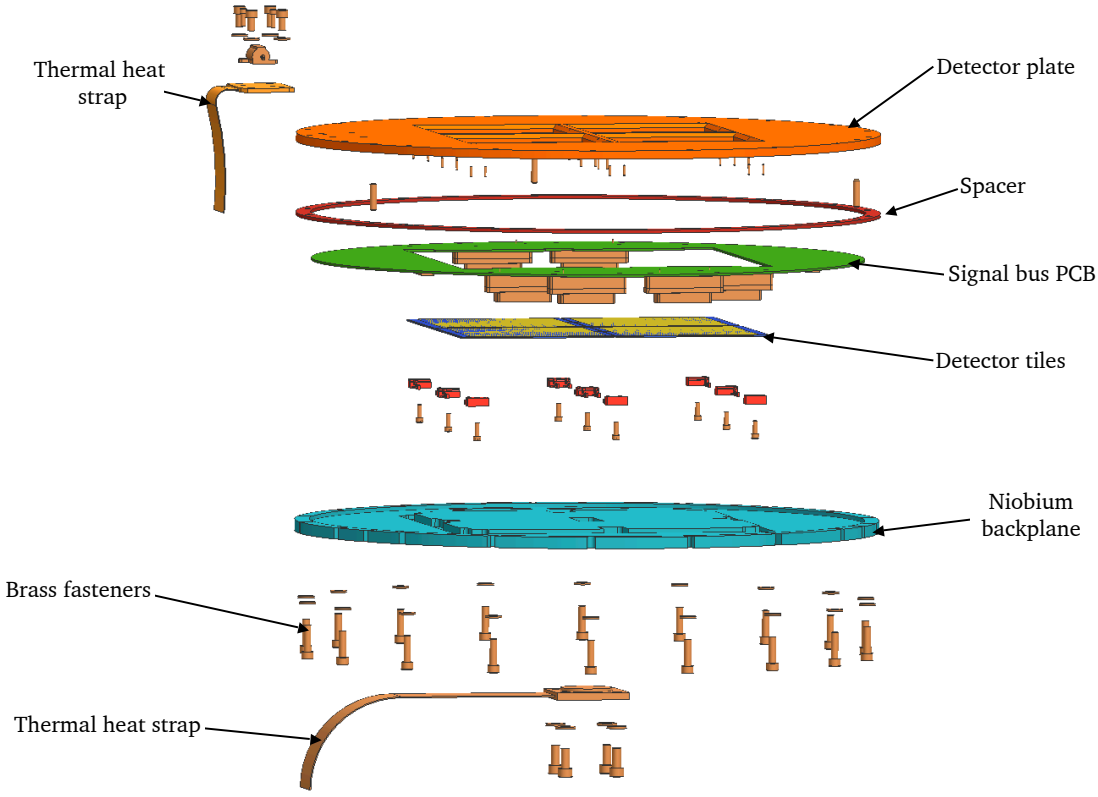


Figure 2.11: Exploded view of an early BICEP2 focal plane. The principal components of the BICEP2 focal plane include a copper detector plate, a signal bus PCB, and a niobium backplate that serves as a magnetic shield and an optical backshort. The SQUID chips are mounted to the signal bus PCB. This exploded view is of an engineering-grade focal plane that was not fielded, but captures the salient features of the design.

niobium backshort. Niobium is a type-II superconductor, and, as a result, flux vortices are pinned in the material. It is thus desirable to cool the shield from the center outwards, expelling flux as it transitions. This minimizes the pinned flux in the plate, and thus reduces the ambient field to which the first and second-stage SQUIDs are exposed. This requires thermal isolation, which is what motivates the separate thermal heat straps for the top and bottom sections of the focal plane, as well as the insulating spacer between the two metal surfaces (Figure 2.11). This spacer also sets the distance for the optical backshort. In BICEP2, Macor washers were used in place of the single continuous spacer illustrated. The niobium backplate was anodized to prevent electrical shorts.

2.5 Cryogenic architecture

The BICEP2 telescope sits within a large cylindrical vacuum vessel. At the center of the vessel is the telescope itself, surrounded by a toroidal liquid helium (LHe) bath with a capacity of 100 L

and a boiloff rate of ∼20 L/day. As the LHe boils off, cold He vapor is forced through a series of small diameter tubes. These are heat sunk to two large cylindrical heat shields that surround the 4 K environment. These vapor-cooled shields remove the need for a liquid nitrogen bath, as was required for BICEP1. Owing to the low ambient pressure, at the South Pole the LHe bath temperature equilibrates to roughly 3.9 K. The inner and outer vapor-cooled shields equilibrate to 40 and 100 K, respectively. The vapor-cooled stages additionally act as heat sinks for the two Teflon filters described in Section 2.2.3.

The telescope is heat sunk to the bottom of the LHe bath by a large copper "camera plate." This is also the heat sink used for the sorption fridge, which is used to cool the focal plane assembly to sub-Kelvin temperatures. When cycled, the fridge generates a substantial amount of heat. As a result, conductivity exceeding 1 W/K is required between the fridge and the He bath, which is achieved through the thick camera plate.

2.5.1 Helium sorption fridge

A three-stage helium sorption fridge provides the cooling power to bring the focal plane assembly to its operating temperature of 280 mK. The fridge comes from Lionel Duband's group at *Commissariat à l'Énergie Atomique*, and similar designs have been used in a number of similar experiments, including Z-SPEC and BICEP1. The closed-cycle fridge consists of three boot-strapped sorption coolers, one of which uses high-purity ^{4}He as the condensate, while the other two use ^{3}He (Figure 2.12). The coldest point of each stage (the evaporator) serves as the heat intercept for the subsequent stage. The ^{4}He stage serves to bring the temperature below the ^{3}He condensation temperature. The intermediate ^{3}He stage serves as a heat intercept, reducing the parasitic load on the ^{3}He ultra-cool stage. Using this boot-strapping scheme, the ^{4}He/^{3}He/^{3}He stages are able to equilibrate to 1.4/0.35/0.23 K, respectively.

The fridge hold time is > 72 hours, requiring the fridge to be re-cycled every 3 days. This is well-matched to the boiloff rate of the LHe tank, which needs to be refilled at similar intervals. The fridge cycle, together with the time taken to reach equilibrium, takes 4-5 hours. Once the focal plane reaches base temperature (∼230 mK), an active thermal feedback loop elevates the focal plane temperature to 280 mK.

2.5.2 Sub-Kelvin architecture

The sub-Kelvin mechanical structure is designed to make the focal plane maximally mechanically rigid, while reducing the parasitic heat load on the fridge. The sub-Kelvin structure consists primarily of two components. The first structure, colloquially known as the "truss," serves to thermally and mechanically isolate the focal plane. The second component consists of a bundle of heat straps

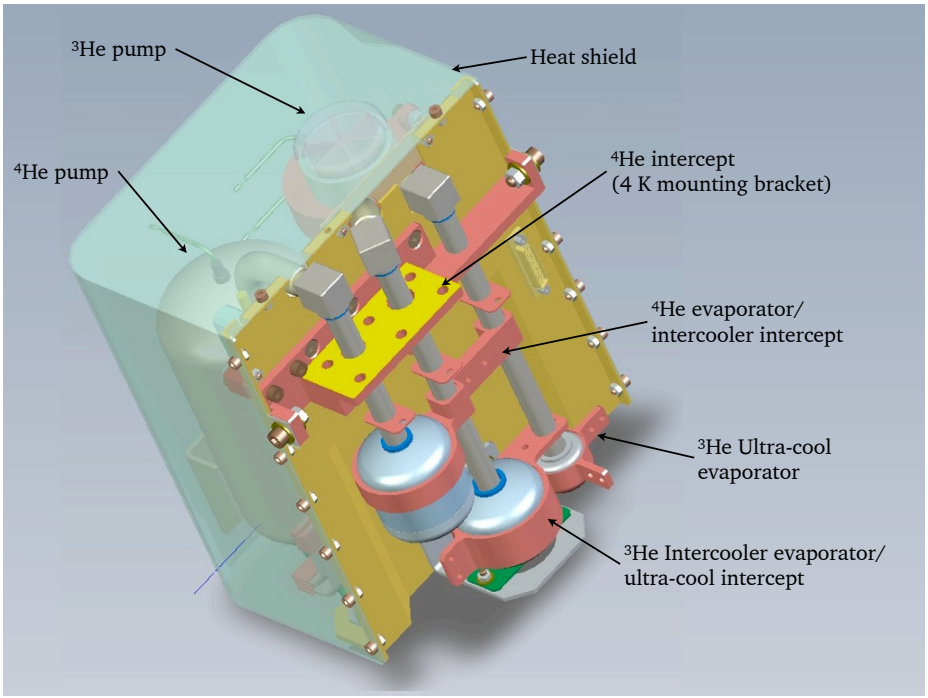


Figure 2.12: The BICEP2 helium sorption fridge. The fridge is heat-sunk to the 4 K bath at the ^{4}He intercept. The ^{4}He evaporator reaches $\sim$ 1.2 K and acts as the ^{3}He intercept. The ^{3}He intercooler evaporator, in turn, acts as the heat intercept for the ultra-cool stage.

Stage	CF conductivity (mW/m)	Total load (μW)
$4 \rightarrow 2$ K	46.5	37.2
$2 \rightarrow 0.35$ K	15.9	8.97
$0.35 \rightarrow 0.25$ K	0.18	0.06

Table 2.2: Parasitic heat loads from the BICEP2 sub-Kelvin carbon fiber (CF) truss structure. Conductivities are calculated using material properties from Runyan and Jones 2008.

used to connect the focal plane and other structures to the evaporator heat sink locations on the fridge.

The sub-Kelvin truss consists of a rigid structure connecting three thermally isolated stages. The design was based on the BICEP1 structure, and modified to improve the mechanical rigidity and thermal isolation. Each stage in the truss is separated by 12 carbon fiber legs, which serve to thermally isolate each subsequent stage (Figure 2.13). The choice of carbon fiber over the Vespel legs used by BICEP1 was largely motivated by the work of Runyan and Jones, which demonstrated the favorable sub-Kelvin properties of carbon fiber rods (Runyan and Jones 2008). The parasitic heat loads from each of the stages from the carbon fiber truss structure are summarized in Table 2.2.

In the first season of BICEP1, it was discovered that mechanical vibrations of the telescope caused temperature spikes on the focal plane. There are two potential sources of thermal excitation at the

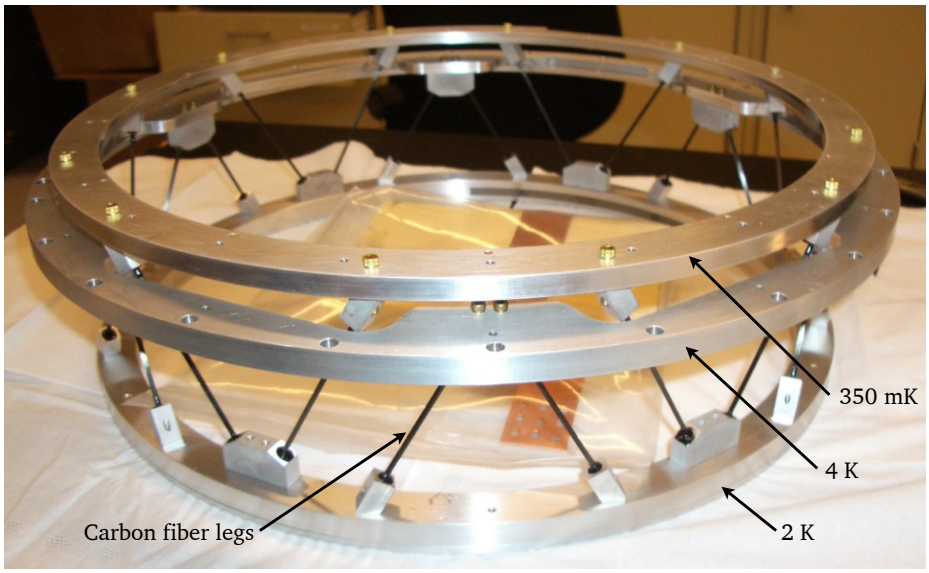


Figure 2.13: Photograph of the sub-Kelvin mechanical support structure. The three concentric rings are heat sunk to 4, 2, and 0.35 K, as labeled. The stages are separated by carbon fiber legs epoxied into aluminum feet. Carbon fiber is chosen for its excellent thermal and mechanical properties.

focal plane: The first is mechanical vibrations that are thermalized at the ultra-cool evaporator point on the fridge. In this scenario, it is predicted that vibrations cause boiling in the superfluid ^{3}He, which, in turn, cause temperature fluctuations that propagate up the thermal heat strap to the focal plane. In the second scenario, mechanical vibrations are thermalized at the focal plane, rather than the fridge. The BICEP2 thermal design was intended to address both potential scenarios.

The BICEP2 focal plane is thermally sunk to the fridge via electron-beam welded flexible heat straps. The straps are constructed from 60 layers of 1 mil oxygen-free high-conductivity (OFHC) copper. The layered copper foil is then welded into copper fixtures at either end, which, in turn, are fastened to the fridge and focal plane. The measured conductivity of the thermal strap is 3 mW/K. To prevent thermal touches of the flexible strap to other thermal stages, the flexible straps are loosely tied to a carbon fiber "spine," preventing the thermal strap from thermally shorting to nearby components.

By design, any mechanical excitation originating at the bottom of the thermal heat strap will be attenuated by the flexible foil, effectively thermalizing any vibration before it reaches the focal plane. For this reason, as well as to attenuate thermal fluctuations at the fridge, we implemented a passive thermal filter, consisting of a large chunk of stainless steel.

The passive thermal filter is placed between the upper section of the heat strap and the focal plane. Because of stainless steel's high heat capacity, the thermal time constant is large (of order seconds). Any high-frequency thermal excitation originating at the fridge or in the heat strap itself is heavily attenuated by this filter. Stainless steel has moderate conductivity, measured to be

0.3 mW/K, roughly an order of magnitude worse than the thermal strap. At equilibrium, however, the temperature gradient is only 10 mK. Since BICEP2 operates 40 mK above the base temperature of the fridge, this thermal gradient is tolerable.

The passive thermal filter is effective at attenuating high to mid-frequency thermal excitations. The thermal transfer function has been measured, achieving > 40 dB attenuation at frequencies above 0.1 Hz (Kaufman et al. 2013a). Low frequency thermal fluctuations are filtered with an active thermal filter by BICEP1-style temperature control modules. The module is a PID-controlled thermometer and resistive heater. Together, the active and passive thermal control achieve a focal plane temperature stability of better than ± 0.2 mK rms during typical telescope scans. The thermal performance of the instrument will be discussed in more detail in later sections.

2.6 Magnetic shielding

A key design consideration for BICEP2 was the magnetic environment within the instrument. There are four magnetically sensitive devices within the instrument; the TES films as superconductors are naturally sensitive to ambient fields. As with all superconductors, an applied field changes the transition temperature of the film. The first and second stage SQUIDs are, by nature, extremely sensitive to applied fields. The superconducting summing coil on the MUX chip is likewise sensitive to ambient fields. Finally, the SQUID series arrays (SSAs) are highly sensitive to magnetic fields.

The magnetic shielding requirements can be calculated from the sensitivity of each of these components to ambient fields. The transition temperature, T_c, of the Ti TES film varies with the applied field strength. According to BCS theory, the transition temperature varies with the applied field according to:

$$\frac{T}{T_c} = \left(1 - \frac{H}{H_0}\right)^{1/2}, \tag{2.10}$$

where H_0 is the critical field strength (measured to be of order $5 - 10$ mT (Lide 2007)), and H is the applied field. We can consider the change in transition temperature to leading order for some change in the applied field, ΔH. Assuming $\Delta H/H_0 \ll 1$:

$$\frac{\Delta T}{\Delta H} \simeq \frac{1}{2}\frac{T_c}{H_0}. \tag{2.11}$$

This yields a field sensitivity of $0.01 - 0.02$ mK/μ. The thermal response of the detectors has been measured to be 0.19 $\mu\mathrm{K_{CMB}}/\mathrm{nK_{TES}}$ (Kaufman et al. 2013b). Assuming a linear response, a change in the transition temperature should have the same effect as a thermal fluctuation of the same magnitude. We thus find that the TES response to magnetic fields is on the order of 260 pT/ $\mu\mathrm{K_{CMB}}$.

To reduce the susceptibility to ambient fields, the BICEP2 SQUIDs provided by NIST are gradiometric, counter-wound into a cloverleaf pattern. This reduces the effective loop area of the SQUID. Even with this improvement, the SQUID response to a uniform DC field is non-zero, in part because a uniform field is naturally distorted by the local magnetic environment. This gives rise to higher-order spatial gradients in the field to which the SQUID is sensitive. The two-dimensionality of the SQUIDs naturally give rise to a direction of maximum sensitivity to magnetic fields. For this reason, magnetic simulations are focused largely on the field amplitude in the direction normal to the plane of the SQUIDs.

We can calculate the sensitivity to magnetic fields at SQ1 as:

$$\left(\frac{\mathcal{R}_{\rm TES} M_{\rm in}}{A_{\rm eff}}\right) \frac{\Delta I}{\Delta T_{\rm CMB}}. \tag{2.12}$$

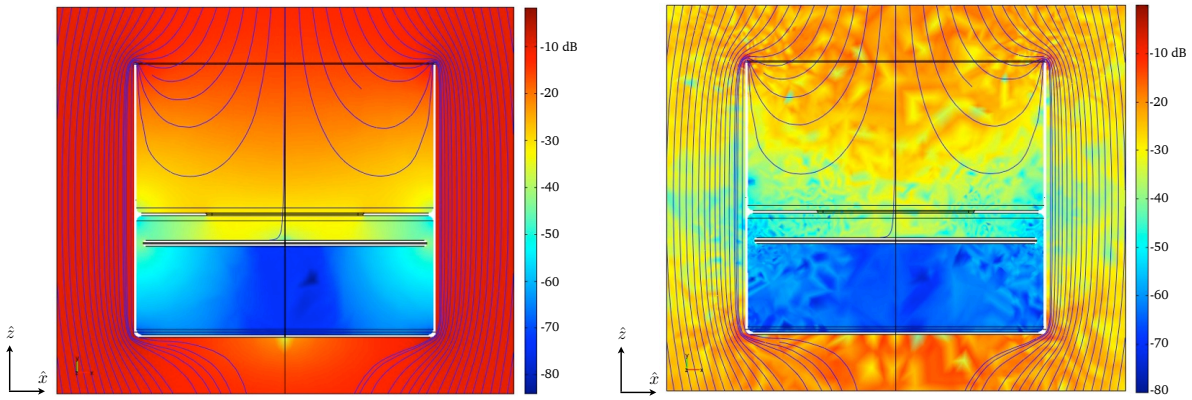
Here $\mathcal{R}_{\rm TES}$ is the TES current responsivity ($\sim 2 \times 10^6$ A/W), $M_{\rm in}$ is the mutual inductance at the SQ1 input coil (475 pH), $A_{\rm eff}$ is the effective area of the SQ1 (measured to be 882 μm^2 in Stiehl et al. 2011), and $\Delta I/\Delta T_{\rm CMB}$ is the CMB power conversion (derived in Appendix A.2). Entering these values into Equation 2.12, we find a field sensitivity of 0.1 pT/μK$_{\rm CMB}$, a factor of 2000 more sensitive than the TESs. Thus, to suppress magnetically induced signal to $< 0.1\mu$K$_{\rm CMB}$, attenuation of -60 dB or larger is required. This, of course, assumes no cancellation or subtraction. In practice, BICEP2's ground subtraction technique (whereby a template of ground-fixed signal is generated and subtracted) reduces common-mode magnetic pickup. Additionally, the pickup will average down over detectors and coverage (since the signal is assumed to be ground-fixed). Our goal is to achieve as much attenuation as possible, keeping in mind the fact that further reduction will be possible in analysis and data averaging.

The second and third stage SQUIDs' magnetic pickup is less concerning than pickup at the first stage. SQ2 is less sensitive, as it has a smaller effective area. Additionally, since the pickup is common-mode within a column, it can be subtracted using either common-mode rejection (pair differencing) or by subtraction of the dark SQUID response. Similarly, while the SSA magnetic sensitivity is high, the response is common to a column. Also, because of their location, the SSAs can be more easily shielded than the first and second stages.

The first line of defense in the BICEP2 magnetic shielding design is a high-permeability nickel-iron alloy (manufactured under the trade name Cryoperm), formed into an open-ended cylinder that surrounds all of the 4 K components. The shield was manufactured by Amuneal.[6] The material is annealed to improve the magnetic susceptibility at cryogenic temperatures.

The focal plane assembly sits within a niobium superconducting "spittoon" that serves as a second stage of magnetic shielding. The geometry was optimized in a series of COMSOL[7] magnetic

[6] http://www.amuneal.com
[7] http://www.comsol.com/



(a) $|\vec{B}|$ (b) $|\vec{B} \cdot \hat{z}|$

Figure 2.14: Attenuation of ambient DC fields within the BICEP2 sub-Kelvin superconducting magnetic shield. Color scale reflects the field density amplitude, while blue lines show magnetic field lines. *Left:* Attenuation of the total field. *Right:* Attenuation of the $\hat{z}$ component of the field.

shielding models performed by Talso Chui at JPL, Marcus Runyan at Caltech, and myself.

The third stage of shielding is the niobium backplane, as pictured in Figure 2.11. The first and second stage SQUIDs lie roughly 1 mm above the niobium backplane. The Meissner effect expels field in the direction normal to the surface, which is the same direction as the direction of maximum sensitivity of the SQUIDs. Because of the boundary conditions imposed by superconductivity, the field normal to the plane, $\vec{B} \cdot \hat{z}$, approaches zero near the surface. With this design, BICEP2 models predict an attenuation factor of -40 dB or larger for the total field amplitude, and -60 to -50 dB of $\vec{B} \cdot \hat{z}$. In later sections, we will discuss the characterization of this magnetic shielding.

The design of the niobium backplane was optimized for magnetic performance. All fastener holes in the niobium are blind, so as to avoid supercurrents and pinned vortices in the material. Similarly, early designs of the backplane included extrusions for the SQUID chips. While seemingly favorable, simulations of this topology predicted large increases in the $\hat{z}$ component of the field, due to distortions caused by the corners of the extrusions. As a result, in later designs, the backplane was made to be as planar as possible. Additionally, a multi-layer sheet of Metglas,[8] a high-permeability metallic foil, was added between the skyward side of the focal plane PCB and the copper detector plate. This additional shielding was not included in our magnetic modeling.

Since the SSAs are physically removed from the focal plane (sitting below in the 4 K "camera tube"), the shielding of these devices was implemented separately. Also, the SSAs are in a self-contained package that is much more easily shielded than the SQUIDs on the focal plane. The SSA module is contained within a niobium box that surrounds the SSA boards on five sides. The niobium can is then covered with multiple layers of Metglas. Because of the physical location of the SSAs, they see little attenuation from the open-ended cylindrical Cryoperm shield. However, the

[8] `http://www.metglas.com`

Figure 2.15: The BICEP2 telescope mount. The blue steel structure enables sub-arcminute pointing in three axes of motion: azimuth, elevation, and boresight rotation.

combination of the Metglas wrap and the niobium box achieve excellent magnetic attenuation, the characterization of which we describe in Section 3.5.

2.7 Telescope mount and drive system

Of all of the subsystems and components shared between BICEP2 and BICEP1, none is more critical than the telescope mount and drive system. All of these components were directly taken from BICEP1, and, as a result, underwent very little change. The telescope mount has three axes of movement: azimuth, elevation, and boresight rotation (referred to as AZ, EL, and DK). The mount at the South Pole was originally manufactured by Vertex for BICEP1, and a second mount was made for calibration measurements. The AZ motion is driven by belt, made possible by the fact that the mount and telescope are never exposed to the Antarctic climate. The EL drive is a contact roller, and the DK drive is gear driven.

2.7.1 Star pointing camera

Pointing is reconstructed using the same optical star pointing camera used by BICEP1. The camera is mounted on the top of the cryostat, aligned with the DK axis. With a 2" resolution, the star pointing camera is used to constrain the telescope pointing model to sufficient precision for BICEP2.

2.7.2 Drive system

The telescope is controlled using GCP (Generic Control Program), originally developed for the CBI experiment. Motion commands are issued via GCP to a Power PMAC (Programmable Multi-Axis Control) controller card (manufactured by Delta Tau Data Systems).[9] The PMAC controls three high-current AE Techron[10] amplifiers, driving three DC brush motors.

The telescope position is established through encoders on each of the three axes. The AZ and EL encoders are angle encoders mounted onto the motors. The DK encoder is a magnetic tape encoder, mounted to the interior rim of the bearing.

2.7.3 Forebaffles and ground shield

To reject spurious reflections and other pickup from the ground, BICEP2 utilizes the same baffling scheme as BICEP1. Directly above the telescope sits a co-moving absorptive forebaffle. The forebaffle is lined with Eccosorb HR-10, and then covered with Volara[11] for weather-proofing. The forebaffle is heated to be just above ambient, to prevent snow accumulation on the interior.

In addition to the co-moving forebaffle, BICEP2 uses BICEP1's ground screen; a large reflective shield that surrounds the telescope. The 8 m diameter ground screen is affixed to the roof of the building, extending 1 m above the window of the telescope. The ground screen is highly effective at reflecting ground pickup, but also well outside of the field of view of the instrument, even at the telescope's elevation limit.

2.8 Observing strategy

The observing strategy for BICEP2 follows directly upon the successful strategy pursued by BICEP1. The sky is mapped in an azimuth-elevation raster at a slew rate of $2.8°/s$. The slew rate has the effect of modulating the science angular scales at a few Hz (at this rate, $\ell = 300$ is modulated at 5 Hz). The entire observing field is covered at four discreet DK angles, oriented at +45, +180, and +225° from the first DK angle. The center of the azimuth scan is held fixed with respect to the ground for 53 minutes at a time. By holding the scan center fixed instead of continuously tracking the field, it is possible to construct and remove a ground-fixed template.

BICEP2 maximizes on-target time by minimizing cryogenic operations, turnarounds, and time spent on calibrations. In this regard, BICEP2 has been enormously successful, achieving a 79% duty cycle during normal operations.

[9] `http://www.deltatau.com`
[10] `http://www.aetechron.com`
[11] `http://www.sekisuivoltek.com/`

2.8.1 Observing site

BICEP2 is located in the Dark Sector Laboratory (DSL) at the Amundsen-Scott South Pole Station, at a latitude and longitude[12] of (89.99° S, 44.65° W). The mount sits on a raised platform on the second floor of DSL, observing the Antarctic sky through an extrusion cut into the roof of the building. The instrument itself is in a short-sleeve lab environment, with only the instrument window and forebaffles exposed to the harsh climate. The telescope is enclosed with an accordion-style vinyl "boot" that moves with the telescope.

The South Pole has long been the preferred observing site for observational cosmologists. A number of other highly successful CMB experiments have been sited at the South Pole, including BICEP1, QUAD, DASI, ACBAR, PYTHON, SPT, and others. It is worth noting that, as of this writing, South Pole experiments can lay claim to i) the most sensitive TT small angular scale measurements to date (SPT, Reichardt et al. 2012), ii) the most sensitive measurements of the EE spectrum to date (QUAD, Brown et al. 2009 and BICEP1, Chiang et al. 2010), and iii) the strongest upper limits on the B-mode amplitude of the CMB to date (BICEP1, Chiang et al. 2010). This highly successful program has been the product of excellent observing conditions and superb logistic support.

With a pressure altitude of $\sim 11,000$ ft and an average temperature of -65 C, the atmosphere above the Antarctic plateau contains minimal precipital water vapor (PWV) and mild, near constant winds. The lack of a diurnal cycle results in long, extended periods with stable observing conditions.

The unique geographic location of the instrument also lends itself to extremely favorable observations of the night sky. Since the observing target never sets below the horizon, observing schedules can be run with an extremely high duty-cycle, interrupted only by cryogenic operations and routine calibrations.

2.8.2 Observing target

The BICEP2 observing target is the Southern Hole, a ~ 800 degree^2 patch of sky accessible from the Southern Hemisphere centered at RA= 0 hr, dec= -57.5 degrees. The observing target covers latitudes distant from the galactic plane, making it exceptionally free from galactic foregrounds. This particular region is the cleanest of its size, with average dust emission averaging 1/100th of the sky median. By sheer coincidence, it is also at a near ideal declination when observed from the South Pole.

BICEP2's observing band at 150 GHz is chosen for three considerations: i) 150 GHz corresponds to the predicted minimum of the sum of galactic synchrotron and dust emission within the observing region (assuming a synchrotron spectral index $k = -3$), ii) as a thermal blackbody at 2.7 K, the CMB

[12] As of 2005. Because of the 10 m/year motion of the polar ice cap at the South Pole, the coordinates of the lab move with respect to the geographic pole.

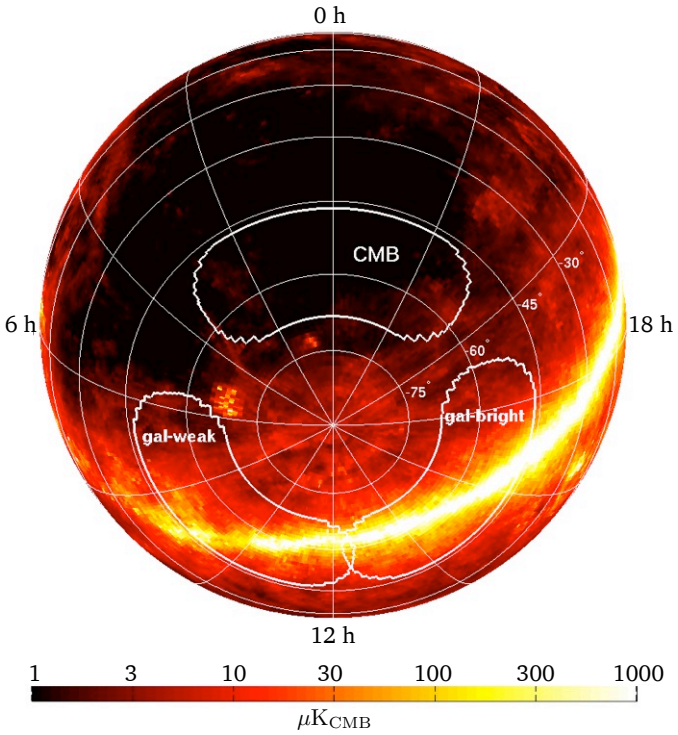


Figure 2.16: The BICEP2 observing field overplotted onto the Finkbeiner, Davis, and Schlegel dust emission model (Finkbeiner et al. 1999). Average dust emission in the BICEP2 target region is over 100 times lower than the sky median. Figure is adapted from Chiang et al. 2010.

peaks at ∼150 GHz, and iii) there is a convenient atmospheric window centered at 150 GHz, between oxygen and water lines at 118 and 183 GHz, respectively. As a single-color instrument, BICEP2 is incapable of distinguishing primordial B-modes from galactic foregrounds, and, as a result, can only place upper limits on the B-mode amplitude of the CMB. In the scenario of a detection of B-mode polarization, further multi-frequency followup will be necessary to distinguish a primordial signature from a foreground.

Chapter 3

Characterization of the BICEP2 telescope

As instrument sensitivity has improved over successive generations of CMB polarization experiments, the requirements for polarization systematics have become increasingly stringent. The threat of instrumental systematics looms large when attempting to measure the B-modes, in part because the analysis relies so heavily on component separation. Small errors in instrument calibration or matching between polarized detectors can leak the (comparatively) very bright temperature fluctuations into polarization. Polarization systematics can be identified both in data analysis and through instrument characterization.

The principal goals of the characterization of the optical, thermal, and magnetic performance of BICEP2 are to: i) characterize the optical and polarization response of the telescope for faithful map reconstruction, and ii) assess and isolate potential sources of instrumental polarization. The latter may be accomplished by generating simulated data with instrument parameters captured from calibration data as inputs. The process of taking measured instrument parameters to constraints on false B-mode polarization will be detailed in Chapter 4.

In this chapter, we summarize our effort to characterize BICEP2. As in the previous chapter, particular attention will be paid to potentially dominant sources of systematics, and to efforts in which I played a substantial role. A non-trivial fraction of my graduate career has been devoted to characterizing the optical response of the telescope. This effort has included three trips to the South Pole, hundreds of hours of data-taking, and thousands of lines of code. Collaborators can view much of the analysis code used to produce the results below in my online working notes, available at `http://bicep.caltech.edu/~rwa/rwa_working_notes/index.html`. In addition to the optical calibration efforts, this chapter will also summarize characterization of the polarization, thermal, and magnetic response of the instrument. We will conclude the chapter with characterization of the instrument performance and noise properties.

3.1 Far-field optical characterization of BICEP2

In this section, we summarize the characterization of the far-field optical response of the instrument in three parts. First, we review the far-field mapping procedure, whereby beam characterization data are acquired, processed, and gathered into maps. This will include a discussion of the beam map construction and beam parameterization. Second, we characterize the beam formed by summing detectors within an orthogonally polarized pair and compare the result with optical models. Third, we assess the pair-difference beam (formed by differencing detectors within a pair).

The primary purpose of characterizing the pair-sum optical response (which we denote as $B_s(\vec{p})$) is to derive an absolute calibration of the BICEP2 temperature maps. As we rely on WMAP for absolute calibration, we re-smooth the WMAP maps with the BICEP2 beam kernel. To leading order the error from uncertainty in $B_s(\vec{p})$ is degenerate with an overall absolute gain factor, but second-order effects can give rise to a slope in the ℓ-space absolute calibration. These second-order effects can often be ignored, even when $B_s(\vec{p})$ is approximated by a simple Gaussian kernel with a matched beam width.

The pair-differenced signal that results from detectors A and B with idealized polarization angles $\psi_A = \psi$ and $\psi_B = \psi + 90$, elevation nod corrected responsivities $\tilde{g}_A$ and $\tilde{g}_B$ (described in Section 4.1.1), and beam functions $B_A(\vec{p})$ and $B_B(\vec{p})$, is:

$$\begin{aligned} d_f(t) =& \frac{1}{2}\int d\vec{p}\,'[\tilde{g}_A B_A(\vec{p}_A(t)-\vec{p}\,') - \tilde{g}_B B_B(\vec{p}_A(t)-\vec{p}\,')]\Theta(\vec{p}\,') \\ &+\frac{1}{2}\int d\vec{p}\,'[\tilde{g}_A B_A(\vec{p}_A(t)-\vec{p}\,') + \tilde{g}_B B_B(\vec{p}_A(t)-\vec{p}\,')](Q(\vec{p}\,')\cos 2\psi(t) + U(\vec{p}\,')\sin 2\psi(t))] + n_f(t). \end{aligned} \tag{3.1}$$

Here, $\vec{p}_i(t)$ is the detector pointing in spherical coordinates for detector i, $B_i(\vec{p})$ is the "as observed" frequency-independent beam function as defined in Equation A.3, and $n_f(t)$ is the pair-difference noise. As in previous sections, Θ, Q, and U are the temperature and polarization anisotropy fields. Detector polarization angles have been assumed to be perfectly orthogonal and the cross-polar response to be zero.

In the event that the beams are perfectly matched ($B_A(\hat{n}) = B_B(\hat{n})$), the term in Equation 3.1 proportional to Θ will vanish. If not, there will be some leakage of the temperature field into the pair-difference time series.

The leakage term in Equation 3.1 is a consequence of the choice to pair-difference before constructing maps. One could construct A and B maps separately, accounting for differences in beam centroids, beam widths, etc., and construct Q and U maps as a final step. For BICEP2, this would come at the cost of an intolerable noise penalty. BICEP2 relies heavily on pair differencing for atmospheric rejection (as well as other common-mode contamination). Alternatively, one could build a

pipeline that simultaneously estimates the atmosphere, the CMB intensity and polarization, ground pickup, focal plane temperature fluctuations, etc., but this is computationally expensive and much more complicated. We thus choose to pair difference "up front," before accounting for differences in the detectors' beams. Characterizing the beam matching between A and B is thus a crucial goal for assessing spurious polarization.

3.1.1 Far-field beam mapping

Far-field beam maps are made possible by the fact that BICEP2 has a relatively close far-field distance, of order ~ 100 m. We can thus make maps of the far-field response using ground-based calibrators, a distinct advantage of small-aperture experiments. Far-field maps have been made in the Caltech synchrotron highbay and at the South Pole station. The Caltech synchrotron highbay measurements were critical for verifying the optical health of the telescope prior to deployment. The calibration data acquired at the South Pole serve as our archival calibration data and will be the focus of our characterization discussion below. To fully capture the rich complexity of the optical response of the telescope, wide field-of-view maps with high signal to noise are required.

Maps of the BICEP2 optical response were made with primarily two source configurations. The first used a broadband amplified noise source. This source consists of a terminating resistor coupled to a chain of amplifiers and frequency doublers, yielding a $\sim 1 \times 10^6$ K microwave source, PIN-switched at 18 Hz. Beam maps were taken with circular and linear polarizers at the output, in both an open-waveguide and horn-coupled configuration. While the maps yielded excellent signal-to-noise, the extreme source brightness led to gain non-linearity in a number of detectors. The source is also highly non-thermal. An example map is shown in Figure 3.1. This map was made with a circularly polarized broadband amplified noise source. This map, with a noise floor at roughly -75 dB below the peak height of the beam, was made in 12 hours of data-taking.

The second source configuration consisted of a chopped reflector, referred to as "thermal source" beam maps. A large-aperture chopper was placed in front of a mirror reflecting the cold South Pole sky, which typically has a zenith temperature of 12 K. The chopper, covered with eccosorb sitting at ambient temperature, modulated the optical signal through the aperture at a rate of 10-18 Hz. This data set proved to yield a much more repeatable and linear main beam response than the microwave noise source. While this came at the sacrifice of signal-to-noise, we still find the map depth sufficient to characterize the main beam performance.

Due to the mechanical restrictions of the mount, the telescope cannot observe ground-based calibrators directly. As a result, a folding flat mirror was required to redirect the telescope's beam to the roof of the Martin A. Pomerantz Observatory (MAPO) at a distance of roughly 200 m (Figure 3.3). Microwave sources were mounted to a mast roughly 10 m above the roof. The use of the folding flat mirror introduces additional free parameters in the pointing model for these beam maps, which

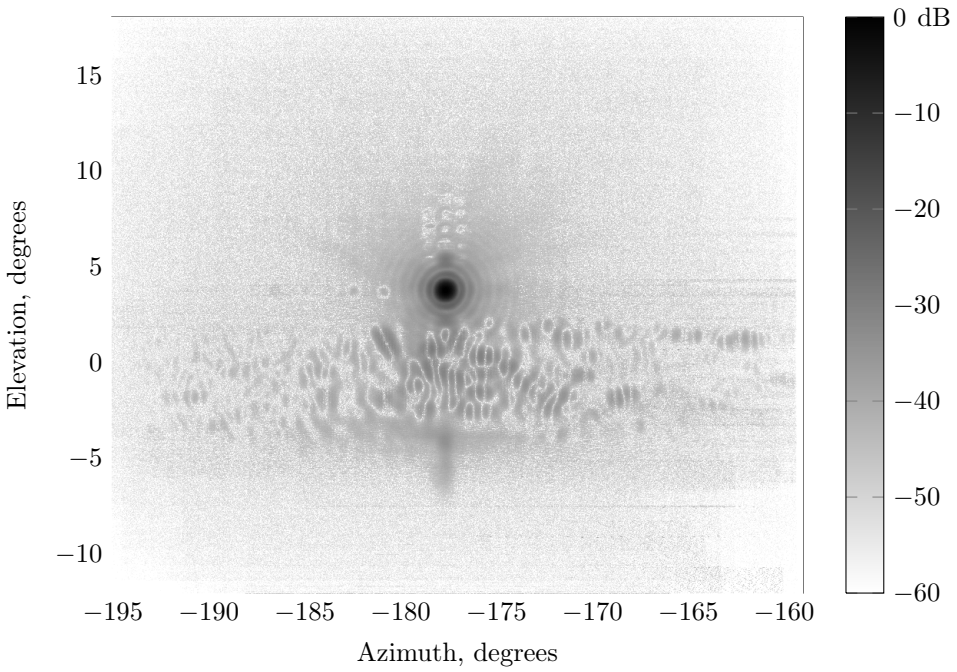


Figure 3.1: Beam map of the BICEP2 far-field response made with a broadband amplified noise source, centered and co-added over all operational channels. Plotted is the logarithmic irradiance in dB. The map, made with the Al TES, spans well over 6 orders of magnitude. Besides the main beam, pickup from the ground is apparent (between $EL = -5$ and 2 deg).

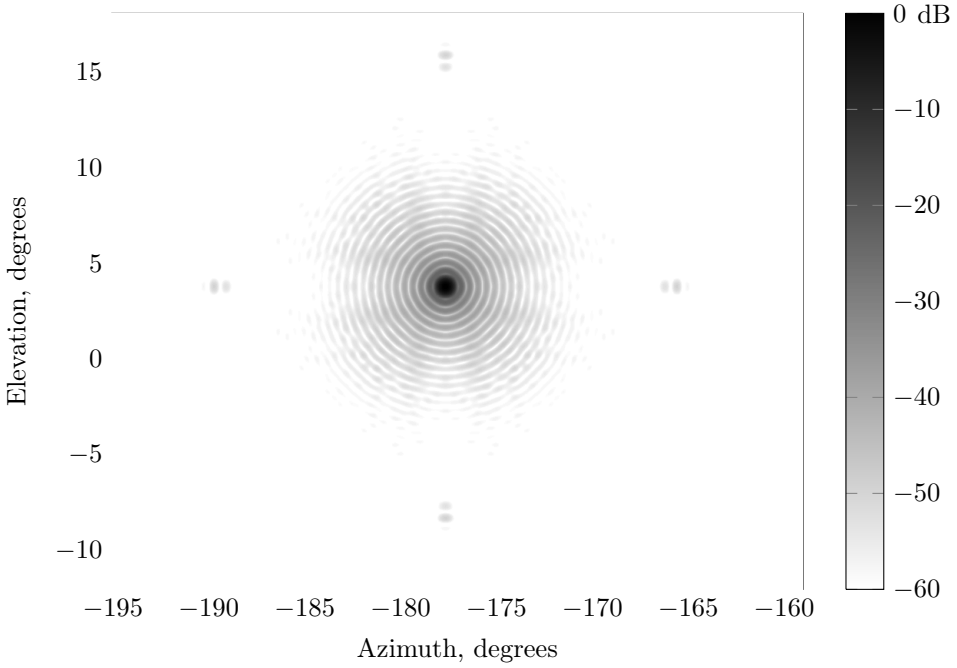


Figure 3.2: Simulated beam map of the BICEP2 far-field response made with Zemax physical optics. The optical model does not include the forebaffle, which has the effect of cutting down large angular response. The measured main beam shape and Airy rings are well-matched by simulations.

we will describe in detail in Appendix C.

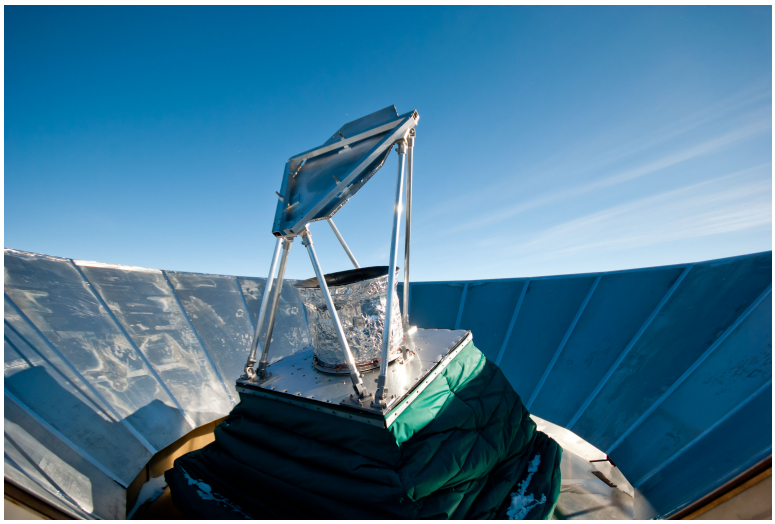


Figure 3.3: BICEP2 at the Dark Sector Laboratory at the South Pole. For beam mapping calibration measurements, a folding flat mirror was mounted above the telescope to redirect the beam to a nearby calibration source (not pictured).

The BICEP2 beam maps are collected in an $AZ - EL$ raster pattern at a single DK angle. In most cases, full rasters were acquired at two or more DK angles. This is a powerful test: beam features that rotate with the telescope can be identified and assessed in this scheme.

When beam mapping, data is collected at a high data rate (relative to the science data taking rate), typically 150 Hz. With the source chopped at $\sim$18 Hz, we acquire many samples over every chop cycle. Azimuthal scans during mapping are driven at scan rate of 2°/s. This ensures that many chops are acquired over the main beam. The telescope encoder data, the detector data and reference chop are acquired synchronously.

The beam map analysis proceeds in three steps: deconvolution, demodulation, then map making. The purpose of the deconvolution step is primarily to account for any group delay that results from a time asymmetric filter that is applied to the data. Typically, the digital Butterworth filter that is applied to science data is disabled. However, in some cases it is desirable to take data with this filter enabled, which has a -3 dB point at 137 Hz. The group delay is accounted for by applying a simple shift to the data stream. A full deconvolution is not necessary, as the beam features are modulated at much lower frequencies.

In the second step, the data is demodulated using a phase-locked square-wave reference chop that is fed from the optical source. The demodulation works essentially by multiplying the reference chop to the detector time series and taking sums over individual chop cycles, resulting in one data point per cycle.

The final, and most complicated step, is the map making itself. This requires a three-dimensional

pointing model that translates raw encoder counts to apparent detector location on the sky. This pointing model is documented in Appendix C. Once all of the per-detector maps have been constructed in ground-fixed coordinates, they can be stacked for higher signal-to-noise, or differenced to form pair-difference maps. In the construction of these maps, the centroid of A and B are assumed to be the same.

3.1.2 Beam parameterization

Before discussing the quantitative characterization of the BICEP2 beams, it is useful to introduce the parameterization used to describe the main beam shape. To begin, we will approximate the BICEP2 beam as an elliptical Gaussian (in later sections, we will investigate deviations from Gaussianity). An elliptical Gaussian has six degrees of freedom: amplitude (1), centroid (2), major and minor axes (2), and ellipse angle (1). These quantities are defined and measured in focal plane coordinates (remaining invariant under DK rotation). This is a non-ideal parameterization for the purposes of polarimetry, particularly because the subsequent pair-difference parameters translate non-trivially to temperature-to-polarization leakage, as we will demonstrate in later sections. Also, the ellipse angle is non-convergent in the presence of noise, making it extremely sensitive to noise fluctuations when the ellipticity is small.

We introduce an alternative parameterization of an elliptical Gaussian beam that is more amenable to temperature-to-polarization leakage estimation. To begin, we first write down the beam function B for the i-th detector. We let $\bar{p}_i$ be the beam centroid of detector i on the sphere, which may be in either ground-fixed coordinates or referenced to the celestial sphere (and thus may vary with time). When we consider a detector pair, we define the common centroid and pair-difference centroid as:

$$\bar{p}_0 = \frac{\bar{p}_A + \bar{p}_B}{2} \tag{3.2}$$

$$\overrightarrow{\delta p} = \frac{\bar{p}_A - \bar{p}_B}{2}. \tag{3.3}$$

We note that both $\bar{p}_0$ and $\overrightarrow{\delta p}$ rotate with DK. The pointing centroid $\bar{p}_i$ is not to be confused with the pointing coordinates relative to the boresight, which are denoted as $\{r, \theta_{\rm pix}\}$.

The beam is measured in an equirectangular projected map, centered on the common beam centroid $\bar{p}_0$. The Cartesian coordinate vector of this map is denoted as $\vec{x}$ and defined as:

$$\vec{x} = [(\vec{p} - \bar{p}_0) \cdot \hat{\imath}]\hat{\imath} + [(\vec{p} - \bar{p}_0) \cdot \hat{\jmath}]\hat{\jmath}. \tag{3.4}$$

Here $\hat{\imath}$ and $\hat{\jmath}$ are defined as:

$$\hat{\imath} = -\sin\theta_0 \sin\gamma_2 \hat{\phi} - \cos\gamma_2 \hat{\theta},\ \hat{\jmath} = \sin\theta_0 \cos\gamma_2 \hat{\phi} - \sin\gamma_2 \hat{\theta}. \tag{3.5}$$

Here ϕ and θ are "physicist" spherical coordinates,[1] θ_0 is the distance of the common detector centroid from zenith, and γ_2 is the angle illustrated in Figure 3.4, calculated as:

$$\gamma_2 = \gamma_1 - \theta_{\rm pix} + 180, \tag{3.6}$$

where γ_1 is calculated via the spherical law of cosines:

$$\gamma_1 = \cos^{-1}\left(\frac{\cos\theta_b - \cos r\cos\theta_0}{\sin r\sin\theta_0}\right). \tag{3.7}$$

Note that by choosing the origin of our coordinate system and the center of our map projection to be equal to the common beam centroid, we minimize distortions from the map projection. The positions of the centroids for A and B are given as displacement vectors from the origin (which is coincident with the common centroid):

$$\overrightarrow{\delta x} = (\overrightarrow{\delta p}\cdot\hat{\imath})\hat{\imath} + (\overrightarrow{\delta p}\cdot\hat{\jmath})\hat{\jmath}. \tag{3.8}$$

The detector centroids for A and B in our Cartesian coordinate frame are given by $+\overrightarrow{\delta x}$ and $-\overrightarrow{\delta x}$, respectively. The beam function for detector i is then given by:

$$B_i(\vec{x}) = \frac{A_i}{\Omega_i}\exp\left[-\frac{1}{2}(\vec{x}-\overrightarrow{\delta x_i})^\top\Sigma_i^{-1}(\vec{x}-\overrightarrow{\delta x_i})\right]. \tag{3.9}$$

Here, A_i gives the integral of B_i, Ω_i is the beam solid angle, and Σ_i is the 2×2 matrix that parameterizes the width and ellipticity (in analogy with a covariance matrix describing a two-dimensional Gaussian distribution):

$$\Sigma_i = \sigma_i^2\begin{bmatrix}(1+p_i) & c_i\\ c_i & (1-p_i)\end{bmatrix}. \tag{3.10}$$

There are thus 6 parameters required to describe a detector's beam: $A, \overrightarrow{\delta x}, \sigma, p$, and c, corresponding to amplitude, displacement from the common centroid, beam width, ellipticity in the 'plus' orientation, and ellipticity in the 'cross' orientation, respectively. Note that the displacement vector $\overrightarrow{\delta x}$ as well as the 'plus' and 'cross' ellipticities are referenced to $\hat{\imath}$ and $\hat{\jmath}$, which are defined with respect to DK. The ellipse parameters p and c can be expressed in terms of the major and minor beam widths and ellipse angle, as in Table 3.1. The beam solid angle is calculated from the determinant of Σ_i as:

$$\Omega_i = 2\pi|\Sigma_i|^{1/2}. \tag{3.11}$$

[1] When referenced to the celestial sphere, spherical coordinates θ and ϕ are related to declination and right ascension, as $\theta = 90 - DEC$ and $\phi = RA$.

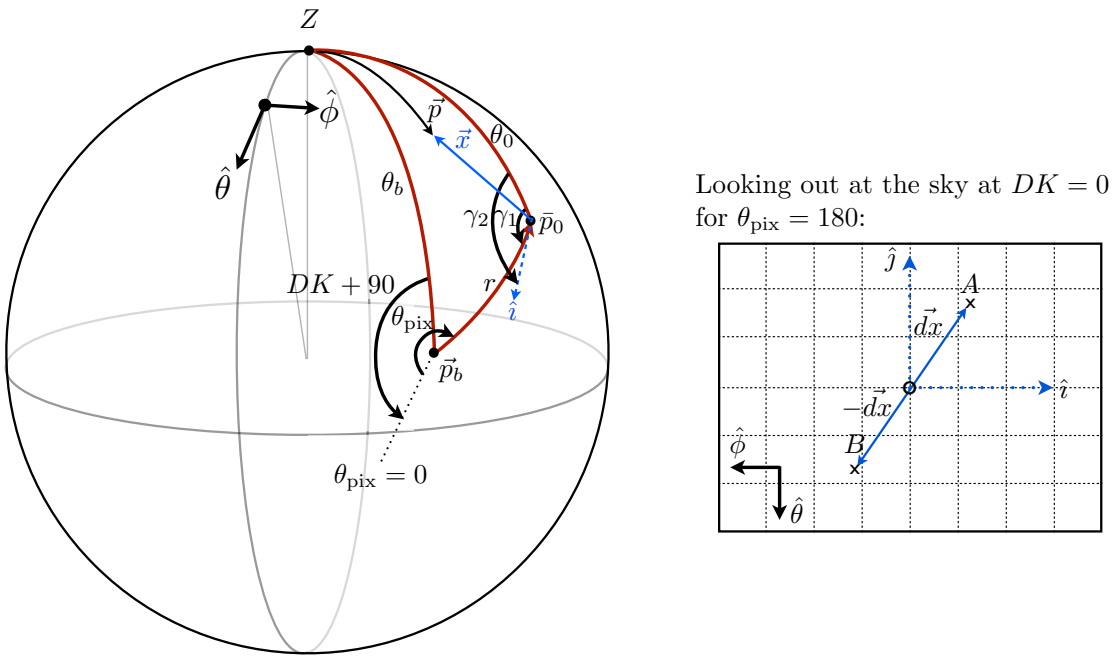


Figure 3.4: Coordinate system used for beam description. *Left:* Spherical coordinate system, which may be referenced to either the ground or to the celestial sphere. On the sphere, the beam is described as a function of variable $\vec{p}$. The location of the common beam centroid is given by $\bar{p}_0$ and may vary with time. We define a Cartesian coordinate system in which to describe a flat-sky approximation of the beam, described as a function of variable $\vec{x}$. *Right:* The Cartesian coordinate system viewed from looking out at the sky at a boresight angle of $DK = 0$. Unit vectors $\hat{\imath}$ and $\hat{\jmath}$ represent an equirectangular projection centered on the common beam centroid, $\bar{p}_0$. Beam centroids for detectors A and B are given as symmetric offsets $\vec{\delta x}$ from the origin.

We can form differential beam quantities by simply differencing the measured beam parameters for A and for B. As we will see in later sections, this results in a convenient and simple parameterization of the resulting leakage from beam mismatch. We find it convenient to construct these differential quantities as symmetric offsets from the pair average. The resulting quantities used to characterize the differential beam are outlined in Table 3.2. The infinitesimal limit of these differential parameters constitute a set of differential "modes", illustrated in 3.5. We implement this parameterization in Section 4.4 to calculate temperature-to-polarization leakage from mismatch in these parameters. As we will show, to leading order, the amplitude of the resulting temperature-to-polarization leakage scales linearly with the parameters in Table 3.2.

3.1.3 Main beam characterization

Our first task is to verify the performance of the beam formed from summing over detectors within a pair. We measure the characteristic beam shape (as estimated by an elliptical Gaussian) and

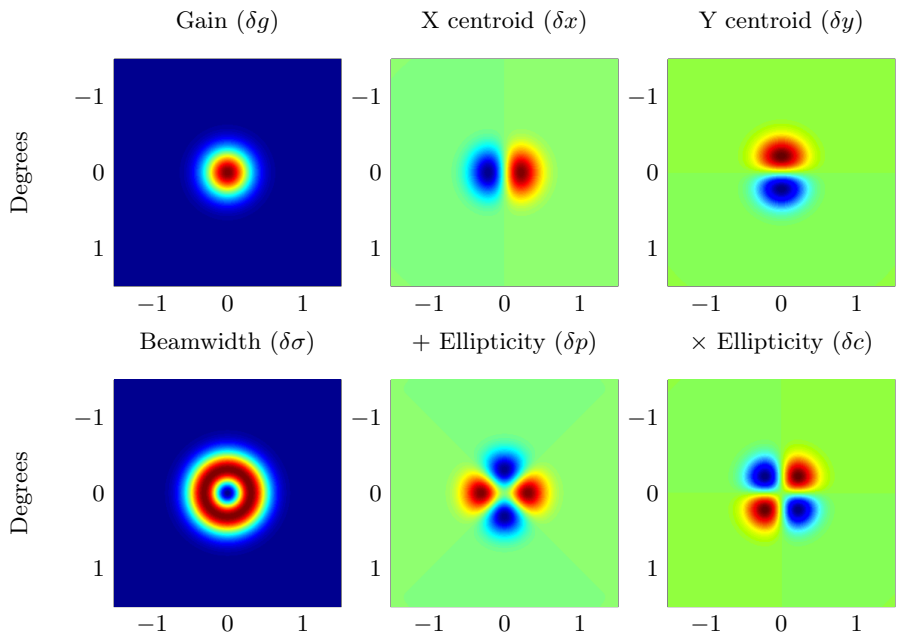


Figure 3.5: Differential beam decomposition. These six modes are defined as infinitesimal perturbations in the a) amplitudes, b) x-position, c) y position, d) beam widths, e) ellipticity in the 'plus' orientation, and f) ellipticity in the 'cross' orientation. To first order, a beam constructed from differencing arbitrary elliptical Gaussians can be approximated as a linear combination of these six modes.

examine consistency with a simulated optical model.

The optical model used for comparison is taken from Zemax physical optics propagation. The model includes all of the active optical elements except for the forebaffle. The diffraction is calculated monochromatically for 150 GHz, giving rise to sharp features in the simulated beam maps that are not present in the real data. A simulated co-added beam is constructed by averaging the results of 8 different detector locations, two in each of the 4 quadrants of the focal plane. Empirically, we find that the beam width of the simulated model needs to be scaled by +5% in order to match the observed beam. This is probably due to a combination of the tapered aperture, as well as a shortcoming of the monochromatic approximation. Comparisons are shown as maps in Figures 3.1 and 3.2. A cross-section of the measured main beam, overplotted with the simulated model, is shown in Figure 3.6. In general, the main beam performance is in strong agreement with the Zemax model. The most immediately obvious differences between the measured and simulated beams are the result of ground pickup, and, at a much lower level, cross-talk. Excess beam response due to cross-talk can be seen in the beam cross-sections illustrated in Figure 3.6. The excess is observed in polarizations A and B at $DK = 0$, when neighboring detectors within a MUX column are co-aligned with the direction of the beam cross-section. We will investigate these cross-talk features more thoroughly in Section 3.2.

Parameter	Symbol	Definition	BICEP2 Measured[1]	Scatter[2]	Error[3]
Beam centroid[4] (arcmin)	$\{r, \theta\}$			1.6'	$< 1.0'$
Beam width (deg)	σ_i	$((\sigma^2_{maj} + \sigma^2_{min})/2)^{1/2}$	0.221	0.006	0.002
Ellipticity (+)	p_i	$\left(\frac{\sigma^2_{maj} - \sigma^2_{min}}{\sigma^2_{maj} + \sigma^2_{min}}\right) \cos 2\theta$	0.013	0.01	0.01
Ellipticity (×)	c_i	$\left(\frac{\sigma^2_{maj} - \sigma^2_{min}}{\sigma^2_{maj} + \sigma^2_{min}}\right) \sin 2\theta$	−0.004	0.03	0.01

[1] Taken as the focal plane mean value.
[2] Taken as the standard deviation over the focal plane.
[3] Taken as the focal plane median per-detector error from repeated measurements.
[4] Beam centroids measured relative to the boresight. The scatter quotes the mean per-detector deviation from a 4 parameter model.

Table 3.1: Measured detector-pair beam parameters for BICEP2. Measured values come from beam maps made using a chopped thermal source in the far field, except for the beam centroids, which are derived from CMB data. Here σ_{maj} and σ_{maj} are the widths along the major and minor axes, respectively, and θ is the rotation angle that diagonalizes Σ, called the "ellipse angle."

The measured detector pair main beam parameters are summarized in Table 3.1. This summarizes the values, scatter, and error on the six parameters that describe the Gaussian beam. The measured values presented are the median per-pair values over the focal plane (rather than a fit to a co-addition over all detectors). The quoted scatter represents the per-pair variability over the focal plane. The quoted error is calculated by taking the median error on the mean from multiple beam maps made under identical observing conditions at multiple DK angles, co-rotated into the same frame. The measured beam centroid is not quoted, as the mean is zero by construction.

The pointing centroids have been measured using both beam map data and CMB data. Centroids from beam map data were derived by performing simple fits and averaging over detector pairs and over multiple measurements. Centroids were derived from CMB data by Walt Ogburn, by cross-correlating per-pair temperature maps with the full co-added map (co-added over all detectors). This cross-correlation was calculated for maps constructed from left-going and right-going scans separately, and also at each of the four observed DK angles, yielding a total of 8 independent measurements. The per-detector scatter is roughly 2 arcmin, yielding an error on the mean of $<$ 1 arcmin. In the end, the CMB data yielded centroids with higher repeatability and less susceptibility to measurement systematics.

3.1.4 Differential beam characterization

As discussed previously, the differential beam characterization is a crucial goal for the success of the BICEP2 experiment, due to potential leakage of temperature anisotropy into polarization. To guide our initial characterization efforts, we used benchmark parameters for differential beam effects taken from Takahashi et al. 2010. These benchmarks were calculated from simulations containing injected leakage from various potential sources of instrumental polarization. For each differential quantity

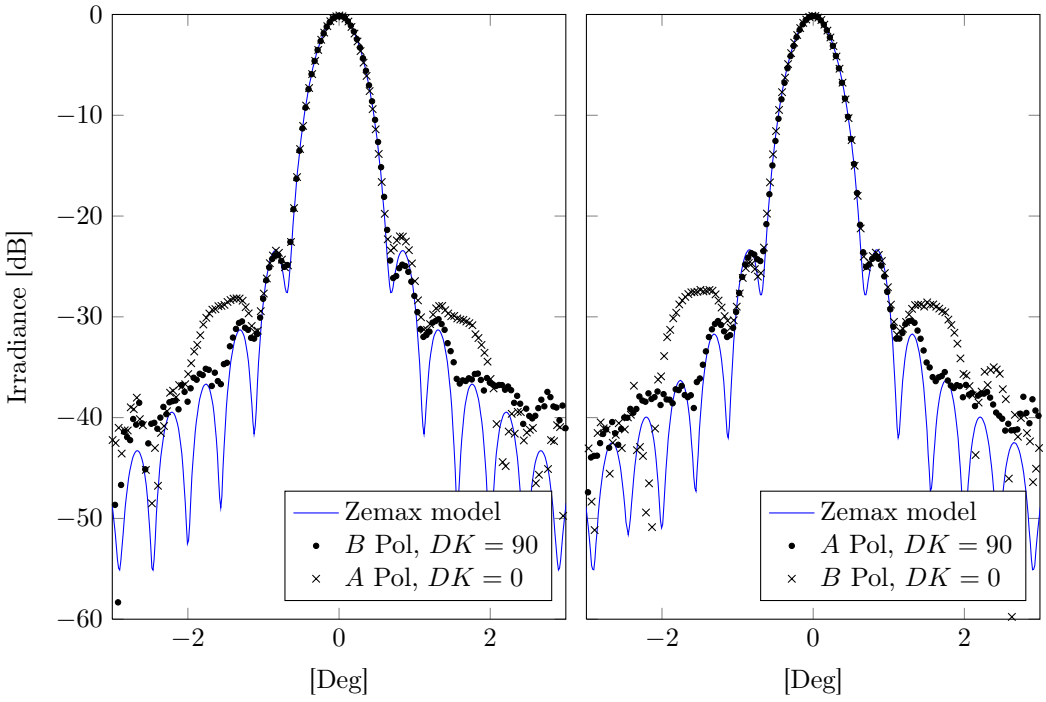


Figure 3.6: Cross-section of BICEP2 beams compared with Zemax physical optics. The cross-section is taken along the scan direction, which is aligned with the horizon. Beam cross sections are plotted with the polarization orientation angles of A and B orthogonal to the cross-section direction (left) and aligned with the cross-section direction (right). The physical optics simulation is monochromatic, giving rise to sharp features in the Airy rings that are not observed in the data. Cross-talk features can be seen in the $DK = 0$ data, in which multiplexer cross-talk within a MUX column gives rise to excess response near the second sidelobe.

(differential pointing, differential beam width, etc.), a benchmark was established by finding the simulated mismatch amplitude that produced excess bandpower equivalent to $r = 0.1$. While these benchmark differential instrument parameters were calculated for BICEP1, they are still applicable to BICEP2 as a consequence of shared scan strategies, similar beam sizes, and parallel analysis pipelines. If anything, using BICEP1 benchmarks to guide BICEP2 will be conservative, as BICEP2 has a higher number of detectors and thus can expect a higher degree of cancellation between detector pairs. That said, the goal of BICEP2 is to eventually achieve sensitivity to $r < 0.1$, requiring tighter benchmarks. We can simply scale the $r = 0.1$ benchmarks calculated in Takahashi et al. by $\sqrt{5}$ to arrive at an approximate BICEP2 benchmark corresponding to $r = 0.02$. For further details on the simulations performed in Takahashi et al., we refer the reader to the original text.

The differential ellipticity benchmark specified in Takahashi et al. is less accurate for BICEP2 because of the method in which it was derived. This benchmark came from BICEP1 simulations that assumed some nominal alignment of the beam major and minor axes based on BICEP1-specific beam map measurements. Nevertheless, we can use this benchmark as an initial estimate to guide

Parameter	Definition	Benchmark[1]	Measured[2]	Scatter[3]	Error[4]
Differential pointing, δx (arcmin)	$\overrightarrow{\delta x} \cdot \hat{\imath}$	0.08'	0.42'	0.56'	0.05'
Differential pointing, δy (arcmin)	$\overrightarrow{\delta x} \cdot \hat{\jmath}$	0.08'	0.41'	0.21'	0.05'
Differential beam width, $\delta\sigma$ (deg)	$(\sigma_A - \sigma_B)/2$	0.002	2×10^{-4}	0.001	6×10^{-4}
Differential ellipticity, δp	$(p_A - p_B)/2$	0.005	0.002	0.013	0.008
Differential ellipticity, δc	$(c_A - c_B)/2$	0.005	0.002	0.007	0.004

[1] Benchmark for the focal plane average, taken as the $r = 0.1$ benchmarks reported in Takahashi et al. 2010 and scaled by $1/\sqrt{(5)}$ to r=0.02.
[2] Taken as the absolute value of the focal plane mean value.
[3] Taken as the standard deviation over the focal plane.
[4] Taken as the focal plane median per-detector error from repeated measurements.

Table 3.2: Summary table of measured differential beam parameters in BICEP2. The per-detector pair parameters are calculated as an inverse-variance weighted combination of a total of 24 beam maps with equal DK coverage. The uncertainty in the focal plane average can be calculated by scaling the quoted per-detector error (Column 6) by the square root of the number of measured detector pairs (194).

our characterization efforts. To arrive at a more precise measure of the predicted bias on r for all of these differential parameters, later stages of BICEP2 analyses will be used to simulate leakage from measured differential quantities.

Differential beam parameters were measured before and after each austral winter observing season, in February 2010, November 2010, November 2011, and November 2012. We measure each of these parameters simply by taking the per-detector fits and differencing within pairs, leading to the results in Table 3.2. Again, these differential parameters have been constructed to be symmetric about the pair-average quantity. The final differential parameters reported in Table 3.2 are weighted averages from 24 separate beammaps, acquired over roughly 170 hours of data taking. The 24 beam maps are split roughly equally between $DK = 0, 90, 180$, and 270. Most of the data weight comes from the November 2011 and November 2012 beam mapping campaigns. By comparing 2011 and 2012 data separately, differential parameters were measured to be constant over time to within measurement uncertainty.

It is apparent from Table 3.2 (as well as from visual inspection of beam maps) that differential pointing (detector-pair centroid mismatch) is by far the most serious differential effect for BICEP2. Of the 206 operational detector pairs on the focal plane, virtually none of them meet the Takahashi et al. criterion for differential pointing. Other differential effects, when averaged over the focal plane, are significantly lower in amplitude. This was clearly measured in the earliest days of BICEP2. Early engineering-grade detectors demonstrated significant offsets in the measured pointing centers. Prior to deployment, multiple generations of detectors were tested in BICEP2, all of which had differential pointing of roughly the same magnitude. While the forensics are still ongoing, it is now clear that the centroid mismatch can be reduced with changes to the detector fabrication process (O'Brient et al. 2012). The differential pointing is somewhat common-mode, as illustrated in Figure 5.2.

Unlike the pointing centers, the beam widths are relatively well-matched between detectors in a polarization pair. Those that are poorly matched are randomly distributed throughout the focal plane (Figure 3.8). Of the 206 operational detector pairs across the focal plane, 180 meet the scaled Takahashi et al. requirements for $r < 0.02$. Future science analyses may consider cutting poorly matched detectors.

Similarly, the differential ellipticity is randomly distributed throughout the focal plane. Characterizing the magnitude of δp and δc separately is necessary for ultimately determining the resulting level of spurious B-mode polarization. However, we can get an initial impression of the beam matching by measuring the magnitude of the total differential ellipticity, $(\delta p^2 + \delta c^2)^{1/2}$, which is plotted in Figure 3.9. Assessing the level of spurious polarization that results from this distribution of coefficients will require simulations in later stages of the analysis.

3.1.5 Polarization efficiency

Using a linearly polarized source in the far field, it is possible to establish upper limits on the polarization efficiency. Far-field maps were acquired with a linearly polarized broadband noise source with an open-ended square waveguide at the output. By modulating the polarization of the source with respect to the focal plane by rotating the telescope about the boresight, we constrained the cross-polar response[2] ϵ to < 0.005. This corresponds to a polarization efficiency of $(1-\epsilon)/(1+\epsilon)$ to be $> 99\%$.

3.2 Un-modeled beam residuals

In this section, we will characterize the anatomy of the measured differential beam that is not captured by a simple elliptical Gaussian model. The ultimate goal of this characterization effort will be to assess the spurious polarization that results from these polarized beam residuals. In practice, this is difficult to accomplish, as it may require convolution of the temperature map with an arbitrarily complicated differential beam. Rather than consider an arbitrary polarized residual, we will instead consider four classes of potential mismatch: near sidelobes, far sidelobes, out-of-band coupling, and cross-talk.

3.2.1 Near sidelobes

In addition to the matching of main beams of two detectors within a polarized pair, we also must consider the matching of the near sidelobes. Here "near sidelobes" refers to optical response outside of the main beam, but within a $0.7-5$ degree radius of the pointing centroid (this is to be distinguished from larger angular scale power, which we call far sidelobes and will review in the next section).

[2]The cross-polar response is the fractional response to a 100% polarized source.

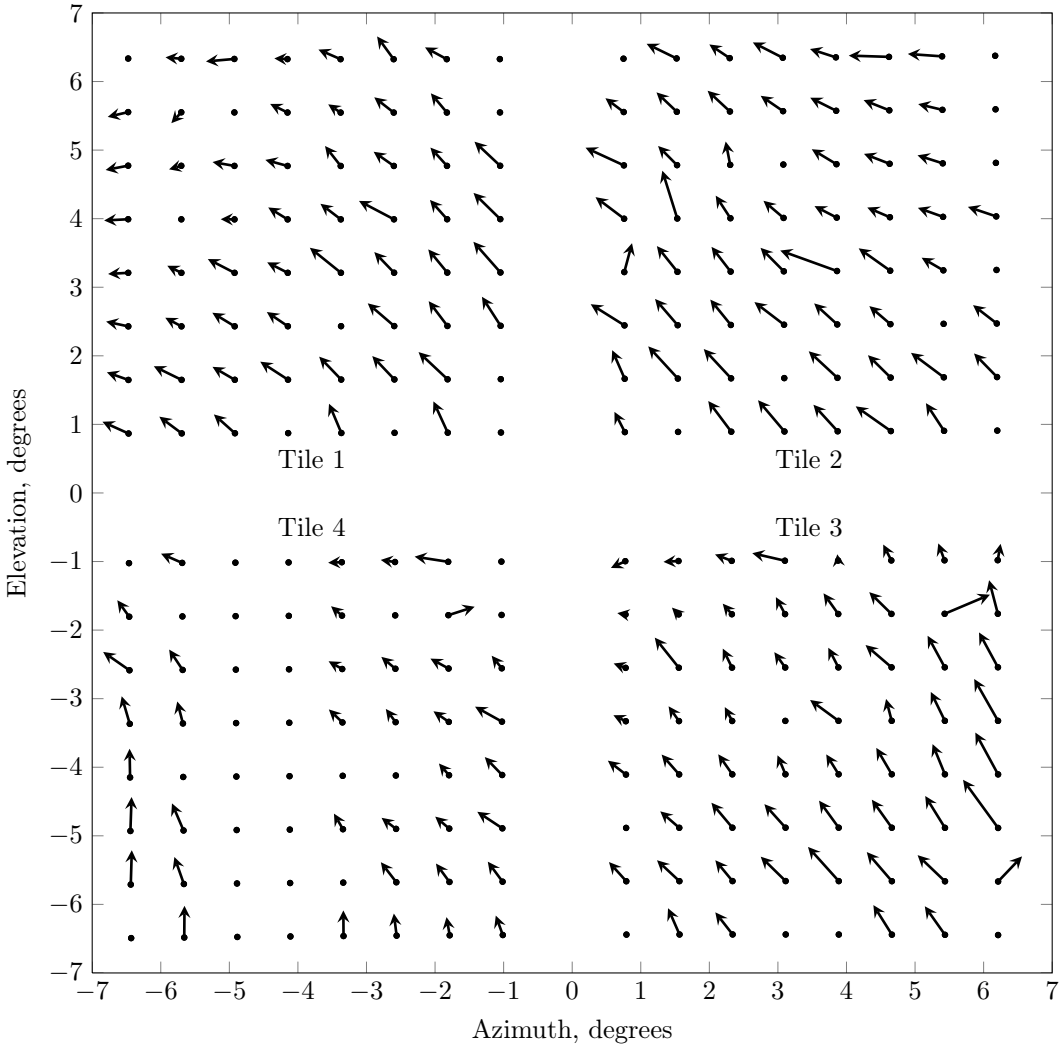


Figure 3.7: Measured differential pointing in BICEP2. Pointing offsets between A and B are illustrated, exaggerated by a factor of 20 for visibility. Some data points are missing for dark or non-operational detectors.

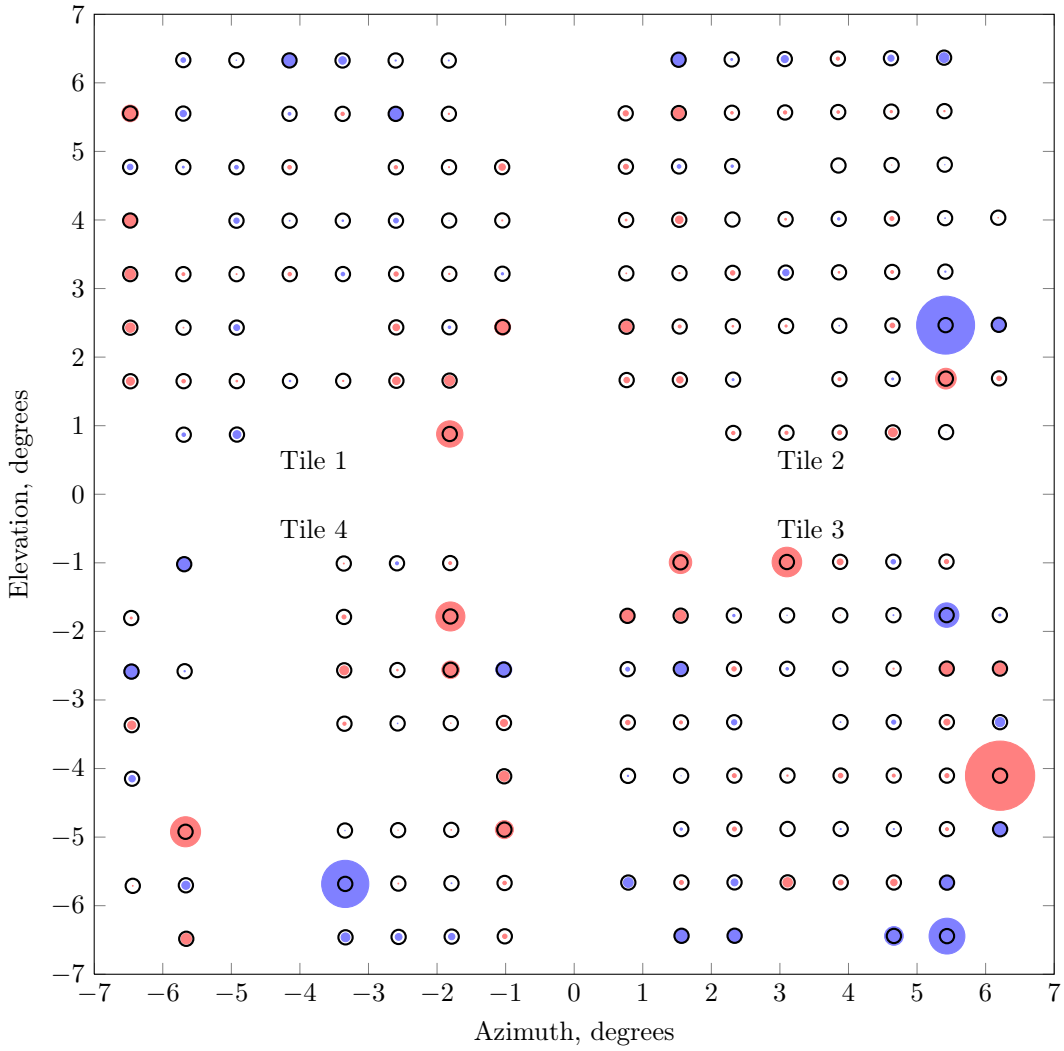


Figure 3.8: Measured differential beam width in BICEP2. The circle diameter is proportional to $\sigma\delta\sigma$ (up to a scale factor), with negative values in red, and positive values in blue. The diameter of the black circles, centered on the pixel centroids, represents the benchmark parameter for differential beam width from Takahashi et al.If the diameter of the colored circle exceeds the black circle, that detector fails the focal plane average benchmark for $r = 0.02$. Unlike the case of differential pointing, the amplitude is largely random across the focal plane. Data points for bad or dark detectors, as well as points for which the measurement uncertainty was extremely high, have been cut. See Figure D.1 for a map of operational pixels.

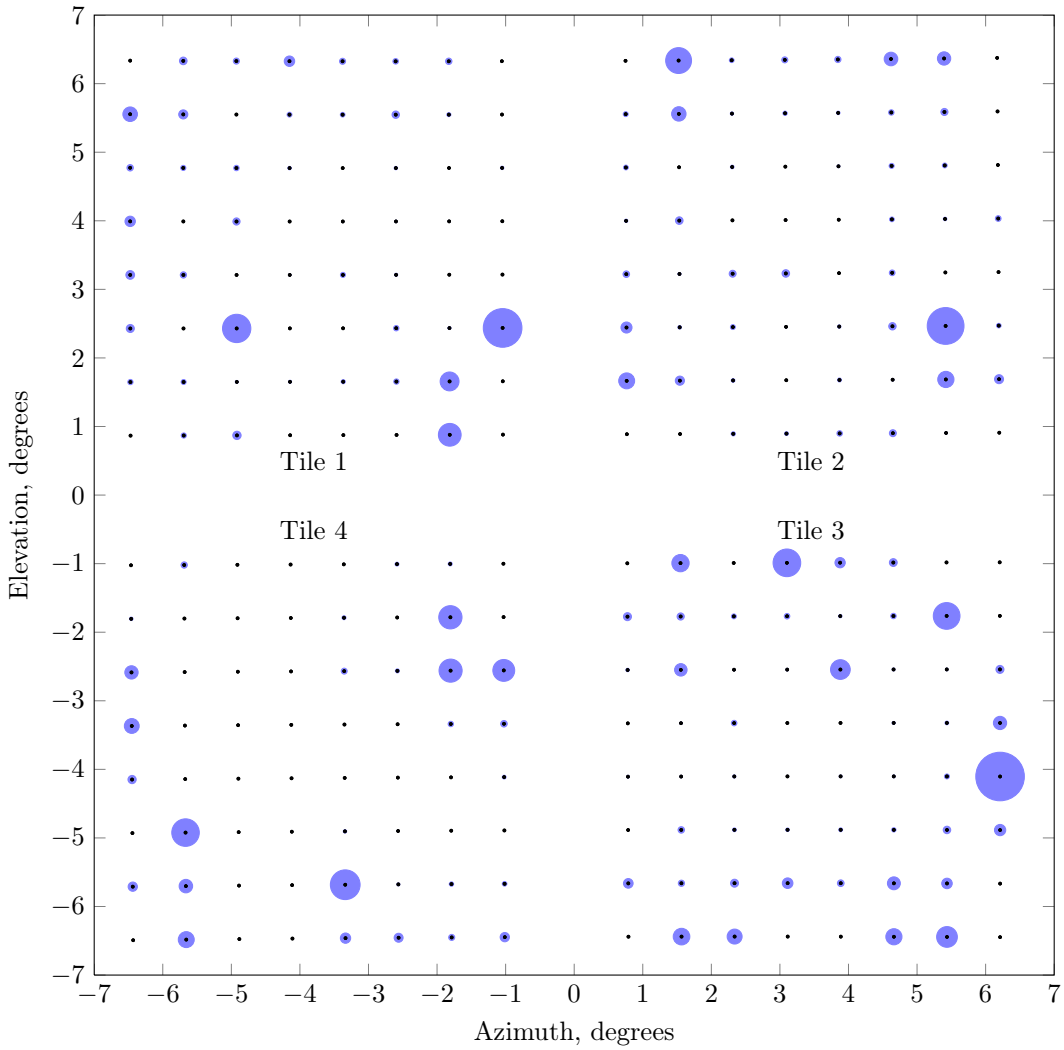


Figure 3.9: Measured differential ellipticity in BICEP2. The blue circle diameters are proportional to $(\delta p^2 + \delta c^2)^{1/2}$ (up to a scale factor), which measures the magnitude of the total differential ellipticity. As in the case of differential beam width, the amplitude is randomly distributed throughout the focal plane, and is consistent with a large population that is well-matched and a smaller population of poorly matched pairs that can be cut if necessary.

Near sidelobes are measured using the far-field beam maps acquired with a circularly polarized broadband microwave noise source (plotted in Figure 3.1). As before, maps are created assuming a common centroid for A and B. In order to isolate sources of polarized sidelobes, maps were co-added over individual detector pairs, tiles, or over the entire focal plane. Investigating these maps with different levels of co-addition enables the identification of mismatch that is unique to detector pairs, or common to a tile or the entire focal plane.

Figure 3.10 illustrates the polarization fraction of the beam out to 2 degrees from the beam center, co-added over tiles. Near the first Airy null, we find that the polarization fraction increases sharply. At larger radii, we can see that the polarization fraction is periodically discontinuous, leaving the impression of radial annuli. The periodicity corresponds to the same periodicity as the Airy rings. While the polarization fraction is high, the integrated power is relatively small; roughly 5% of the integrated optical power lies outside of the first Airy null. The source of the polarization of these sidelobes is not currently understood. In amplitude, the primary source of near sidelobe mismatch is a polarized first Airy ring, which can be seen in Figure 3.11 at roughly 0.8 degrees from the beam centroid. This is apparent in beam maps as well as in this azimuthally averaged radial profile. Further out in radius from $1-5$ degrees, the azimuthally averaged polarized fraction of the sidelobe is roughly $1-10\%$.

Because of the complex nature of the near sidelobe pattern, predicting the resulting level of temperature-to-polarization leakage in power spectra is non-trivial, though it is not a completely intractable problem. Efforts to construct models of the near sidelobe patterns as inputs to simulations are ongoing, and may ultimately be a necessary test to assess the potential polarization leakage from mismatched sidelobes.

3.2.2 Far sidelobes

Looking beyond the near sidelobes, we also consider potential spurious polarization from far sidelobes, that is to say, optical response well away (> 5 degrees) from the main beam centroid. The distinction between near and far sidelobes may seem arbitrary, and, indeed, 5 degrees does not represent a firm threshold. What distinguishes these two residuals is the leaked signal. In the case of near sidelobes, differential optical response will tend to leak the nearby CMB temperature field into polarization. In the case of far sidelobes, one must consider potential differential optical response to much brighter, and in some cases highly polarized, sources. This could potentially include optical coupling to galactic emission, the ground shield, the sun or moon, or directional RFI. Polarization contamination can thus pose a serious threat, even if the relative optical response is quite low. It is for this reason that wide-angle, high signal-to-noise sidelobe maps are necessary.

Far sidelobe data were acquired in the 2011 and 2012 austral summers. In both instances, maps were acquired by observing a $\sim 1 \times 10^6$ K microwave noise source at a range of 10 m. The telescope

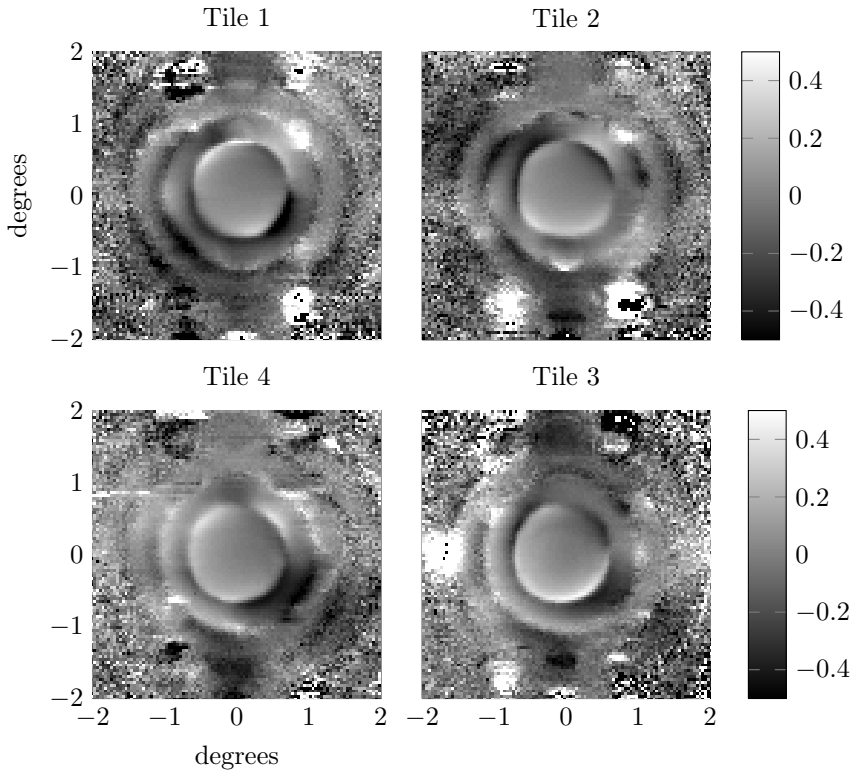


Figure 3.10: Polarized fraction of near sidelobes. Outside of the main beam, the near sidelobes are highly polarized (up to 50%), but contain little integrated power. Sharp discontinuities in the polarized fraction occur near Airy nulls. The polarization of the near sidelobes is not predicted by optical models and is not currently understood.

was rastered in a pattern covering its full range of motion. Beginning at vertical ($EL = 90$), the telescope slewed through 360 degrees in AZ at a fixed DK angle. The elevation was then lowered in increments of 0.5 degrees down to the elevation limit, repeating the azimuthal scan at every step. To cover as much of the field-of-view as possible, the DK was rotated by 180 degrees at the elevation minimum. The scan was continued at interstitial elevation steps back up to vertical, repeating the same azimuthal scan at each step.

These far sidelobe maps were acquired with many combinations of instrument and source configurations. Some of these different source configurations were necessary because the dynamic range required of the measurement (roughly 10 orders of magnitude) exceeds the dynamic range of the instrument, even on the Al transition. With the microwave source full-open, the detectors came unlocked when within a few degrees of the main beam. It was thus necessary to "stitch" two or more measurements together at different source amplitudes. These multiple measurements make it possible to probe deeply for far sidelobe response, and to measure the amplitude relative to the main beam. Maps were acquired with and without the co-moving forebaffle (described in Section 2.7.3) to assess the performance of the baffling scheme.

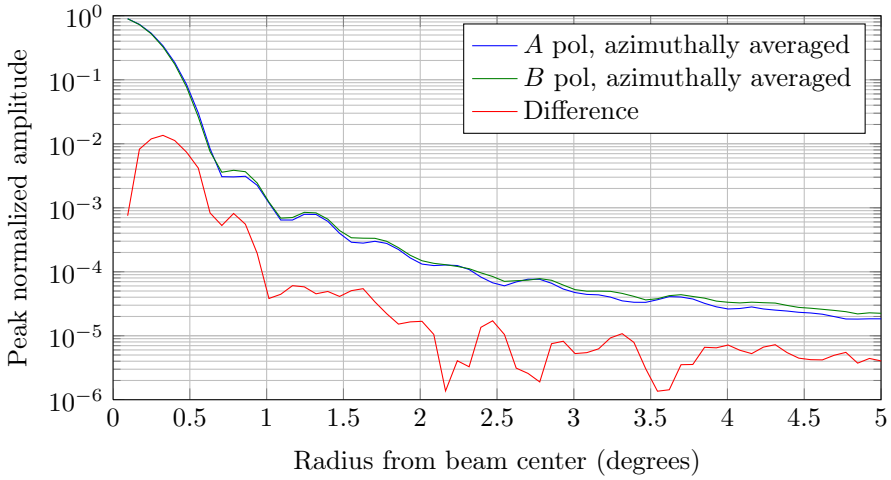


Figure 3.11: Measured near sidelobes, azimuthally averaged about the beam center co-added over all A and B detectors. The difference is plotted in red.

A further complication of these measurements is the polarization of the source configuration. As the telescope rotates in AZ, the polarization state of the source projected onto the polarization axes of the detectors rotates as well. For this reason, maps were acquired using two orthogonal linear polarizations to fully reconstruct the co-polar and cross-polar response for all detectors.

The map construction proceeds in much the same way as described in Section 3.1.1. We employ a similar pointing model, as described in Appendix C, to account for parallax effects, but with no folding flat mirror. Using the reconstructed pointing information, maps are binned in ground-fixed AZ and EL.

Far sidelobe maps are plotted in Figure 3.12. As with previous maps, these have been co-added over all A and B detectors in the focal plane separately. The annular high noise stripe in the figure is the result of the bright source driving the detectors out of lock. When this happens, the detectors continuously ramp through the full range of the feedback loop. These channels are then cut, resulting in a strip with much less integration depth than the rest of the map. Similarly, the beam center is not accurately measured in these maps due to the same dynamic range issue.

Several obvious features are present in this data: first, a square-ish pattern can be seen at a radius of roughly 15 degrees from the main beam. Secondary reflections within the optical chain of the instrument are responsible for this excess pickup. These secondary reflections (anthropomorphically referred to as "little buddies") are an optical artifact caused by reflections off the focal plane. They manifest themselves as ghost images of the main beam mirrored through the optical axis. This is what gives rise to the square pattern when co-added over detectors: Since the detectors are arranged in a square pattern across the focal plane, so too will be the "little buddies." Similar optical ghosts were found in BICEP1 and measured to have an integrated amplitude of -22 dB (Chiang 2008). Due

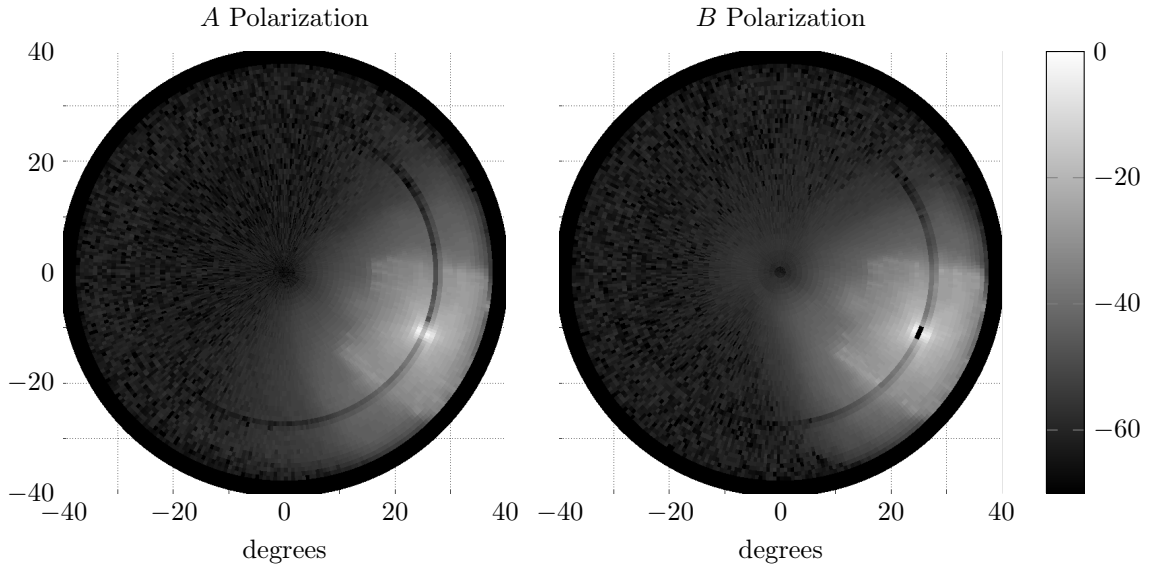


Figure 3.12: Measured far sidelobes co-added over A polarized detectors (left) and B polarized detectors (right). The maps are presented in a gnomonic map projection, with the topocentric zenith at $r = 0$. $AZ = 0$ runs along the positive x axis. Prominent features in these maps include: i) an annular stripe at a radius of ~ 28 degrees, which is a data-taking artifact of losing lock on the bright source, ii) a square-shaped diffuse pattern around the main beam from optical ghost images (or "little buddies"), and iii) diffuse pickup at wide angles that results from scattering within the optics, as well as secondary reflections within the telescope.

to improved anti-reflection coating, the integrated power in the ghost beam has been reduced to < -30 dB, relative to the detector's main beam. Crucially, for both BICEP1 and BICEP2, these ghosts were measured to be largely unpolarized, and thus not a significant contributor to spurious polarization. Because of the scan pattern of the far sidelobe measurement (repeated at 180 degrees relative to the original scan), the amplitude of this feature appears artificially doubled in Figure 3.12.

The second obvious feature in these maps is diffuse pickup between roughly 15 and 30 degrees from the beam center. At larger radii from the main beam, the optical response drops considerably. This is the co-moving forebaffle at work: wide angle response is largely rejected, and within measurement uncertainty consistent with zero. The diffuse pickup that is within the solid angle defined by the forebaffle has contributions from scattering from various optical components as well as secondary reflections within the telescope. The diffuse pickup from secondary reflections within the optics chain is seen as a sharp focus in the aperture plane, which has been measured with near-field beam maps.

3.2.3 Out-of-band coupling

Early generations of BICEP2 detectors demonstrated substantial ($\sim 2\%$) out-of-band optical coupling. This was measured by placing a bright thermal source with a 3 cm aperture directly above the window of the instrument and blocking all other light paths into the instrument. In front of this aperture we attached a variety of spectral filters, both above and below the detector's design spectral passband. We then examined the response of both light and dark detectors.

Dark detectors (having a severed connection to the antenna and summing tree) exhibited significant optical response, particularly at frequencies above the cutoff of the in-line filter. The pickup was largely unpolarized and non-directional, and is now believed to have been caused by direct optical coupling to the TES island. The summing tree "trunk" and the TES bias lines combine to form a dipole antenna. These lines can be seen in Figure 2.6 as the microstrip lines running along the suspending legs to the right and left, terminating on the TES island. Microwave power is believed to couple directly to these transmission lines and terminate on the island. As a result, the microwave power bypasses the in-line filter and thus extends well beyond the design spectral passband.

The direct island coupling was substantially suppressed by reducing the cutout area of the ground shield around the TES island. To further mitigate any potential out-of-band pickup, a low-pass metal-mesh filter was added to the optical stack (described in Section 2.2.3). Even with these improvements, a small level of dark pixel coupling has been measured. Using a chopped thermal source directly above the instrument and a high-pass thick grille filter (with a nominal cutoff of 220 GHz), out-of-band pickup has been measured at an amplitude of 0.3%, relative to the in-band response to the same source (and accounting for the filling-factor of the grille). Dark detectors see roughly the same amplitude of pickup, consistent with some residual direct coupling to the TES island.

The in-band dark pixel optical coupling is roughly the same amplitude, but highly variable throughout the focal plane (between $0.2 - 0.5\%$). By performing the same optical tests with the dark detectors off transition, we rule out any substantial contribution to the dark response from cross-talk. For light detectors, we expect that the total (including both in-band and out-of-band) direct-island coupling is sub-percent.

The spatial distribution of the out-of-band optical response was tested by making far-field beam maps (of the same variety described in 3.1.1), but with a high-pass filtered source with a cutoff above the passband of the detectors. We found that the out-of-band coupling did not form a sharp focus in the far field.

3.2.4 Cross-talk

The multiplexing (or "MUX") architecture of the electronics readout results in a variety of different cross-talk mechanisms between detectors. These coupling mechanisms are described and characterized in detail in de Korte et al. 2003. The first substantial cross-talk mechanism is through inductive coupling between the first-stage SQUIDs and neighboring input inductors. The cross-talk is thus symmetric between channels and independent of the MUX rate and the order in which the detectors are addressed. An additional potential source of cross-talk is "settling time" cross-talk, whereby a detector's signal extends beyond the time window allocated for the channel. Particularly when the previous channel sees a large signal, the feedback loop's recovery time may overlap with the readout of the subsequent detector. In contrast to inductive cross-talk, this cross-talk is asymmetric; upstream channels leak into downstream channels, but not vise-versa. Additionally, settling time cross-talk is sensitive to the details of the multiplexing configuration, and can be mitigated by decreasing the multiplexing rate (thereby increasing the dwell time for each channel). This source of cross-talk is particularly severe for channels that have lost lock entirely. A third source of cross-talk results from multiple first-stage SQUIDs being turned on simultaneously. This is not an unavoidable mechanism, but rather a result of suboptimal tuning of the first stage SQUIDs or electrical shorts on the focal plane. Through beam mapping, we identified 9 such readout channels that appear to optically respond to multiple pixels on the focal plane. These have been cut from analysis.

These and other sources of cross-talk can be assessed in far-field beam maps. By "lighting up" a single detector (with a polarized source) or a detector pair (with an unpolarized source), cross-talk artifacts can be seen in neighboring detectors. Inductive cross-talk has been measured to be on average 0.3% amplitude (-25 dB), consistent with the amplitudes reported in de Korte et al.[3] Because inductive cross-talk is symmetric, A and B detectors within a pair will cross-talk into one another, appearing exactly as a reduced polarization efficiency. We are thus restricted to directly measuring the inductive cross-talk amplitude of only one neighbor (the MUX neighbor that is displaced on the sky). Settling time cross-talk was measured to be substantially lower, with a median amplitude of 0.037% (-34 dB).

Other cross-talk mechanisms have been observed in BICEP2, but most without clear optical counterparts in far-field beam maps. In a few MUX columns, compact noisy spots appear in far-field data when channels within a MUX column come unlocked. When detectors come unlocked, the feedback loop continuously ramps across the full range of the DAC, so even small levels of inter-column cross-talk are made evident. These artifacts are a consequence of observing a very bright source and should not impact typical science data acquisition.

[3]Note that the amplitudes reported here are $10\log_{10}$, while de Korte et al. report $20\log_{10}$.

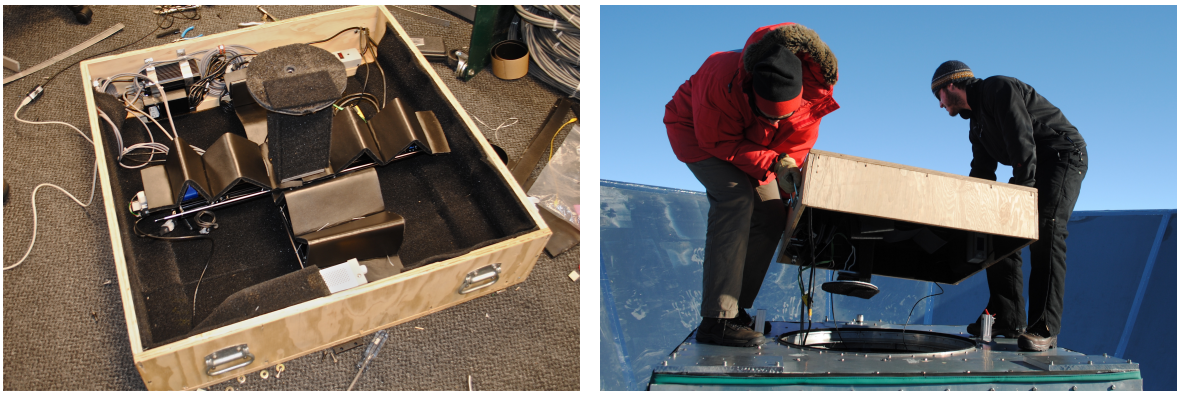


Figure 3.13: Near-field beam mapper at the South Pole. *Left:* The mapper nearing completion in the lab prior to installation. A chopped thermal source sits behind the circular aperture plate that is mounted to an $x - y$ linear translation stage. *Right:* Installation of the mapper atop the "boot" of the telescope by John Kovac (left) and myself (right).

3.3 Near-field optical response

A key diagnostic of the instrument's optical performance during both the engineering and science-grade phases of the experiment was the near-field optical response of the instrument. While far-field beam maps are primarily sensitive to the amplitude distribution of the electric field in the focal plane, maps made in the aperture plane of the instrument are primarily sensitive to the phase of the electric field in the focal plane. As a result, near-field maps can serve as a probe of phase gradients within the phased-array antennas. Near-field maps can also serve as a probe of secondary reflections that focus near the aperture plane as well as vignetting within the telescope.

Near-field maps were made with an $x - y$ linear translation stage attached to a large blackened enclosure and placed directly above the window of the telescope. The source used for mapping was a ceramic heater covered with high-temperature silicone loaded with carbon lamp-black. The blackened surface was measured to be less than 3% reflective and highly absorptive. The source was chopped by a simple wheel chopper running near 18 Hz. A machined and blackened aperture plate with a roughly 1 cm diameter aperture was placed in front of the source and chopper, just a few finger-widths from the window. A large ($\sim$ 40 cm) slew made it possible to map the response over the entire telescope aperture area.

Near-field maps were acquired during two successive summer seasons at the South Pole. (Pictures of the mapper and the installation can be seen in Figure 3.13.) Maps were acquired in a step-dwell raster pattern at several DK angles. Like the far-field beam maps, the data were acquired on the Al transition. Example maps are plotted in Figure 3.14 for A, B, and $(A - B)$. The sharp bright feature in the bottom right quadrant of the maps is a secondary reflection from the 4 K spectral filters refocused into the aperture plane. This spot contains less than 0.1% of the integrated power of the main beam. Moreover, since it forms a sharp focus in the aperture plane, it must be broadly

and diffusely coupled to the sky in the far field.

Near-field maps acquired both before and after deployment revealed two non-idealities in the optical performance of the instrument. The first is a "beam-steering" effect, where the main beam of the detectors appears steered into the aperture by $5 - 10$ degrees, substantially more than is predicted by any physical optics model. This beam effect was described in detail in Aikin et al. 2010. This steering is readily apparent to the eye, and was, for a few pixels, so severe that the main beam was steered completely off of the aperture and into blackened surfaces. This impacts not only optical throughput, but also can potentially introduce beam distortion caused by the asymmetric and aggressive illumination of the Lyot stop. In general, this steering was not particularly well-matched between detectors in a polarized pair, resulting in mismatched beam shapes in the far field. The BICEP2 focal plane contains roughly 50 detector pairs that are steered at some level, while only roughly a dozen suffer from the severe steering illustrated in Figure 3.14.

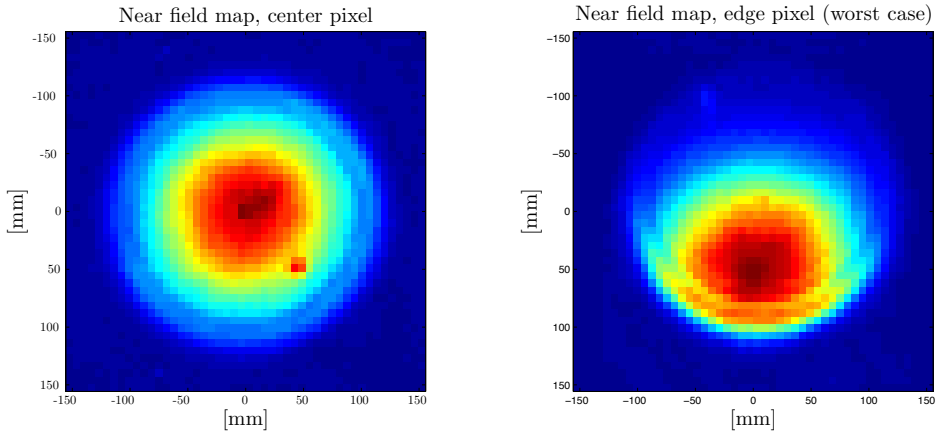


Figure 3.14: *Left:* Near-field map of a pixel near the center of one of the detector tiles in BICEP2, with typical beam performance. *Right:* A highly steered detector (worst case) for a pixel near the edge of the detector tile. Plots are made on a linear color scale, both peak-normalized to unity. Phase gradients in the phased-array antenna were demonstrated to be responsible for the steering. Adapted from Aikin et al. 2010.

We found a great deal of tile-to-tile variation in the level of the beam steering. Additionally, the worst offenders were found to be consistently along the top and bottom tile edges. A large number of possible culprits for the steering were considered, including distortion from the detector plate holding the detectors in place and the optical backshort. Testing of devices fabricated after the BICEP2 deployment demonstrated a dramatic sensitivity of the severity of the steering to the details of the device fabrication process. It is now believed that the device uniformity was being degraded toward the edge of the tile, leading to phase gradients in the microstrip summing tree. While this has been improved for later generations of detectors, BICEP2 may require simply cutting the pixels that suffer most severely from this beam steering effect.

The second non-ideality revealed by the aperture plane maps is a near-constant A/B beam

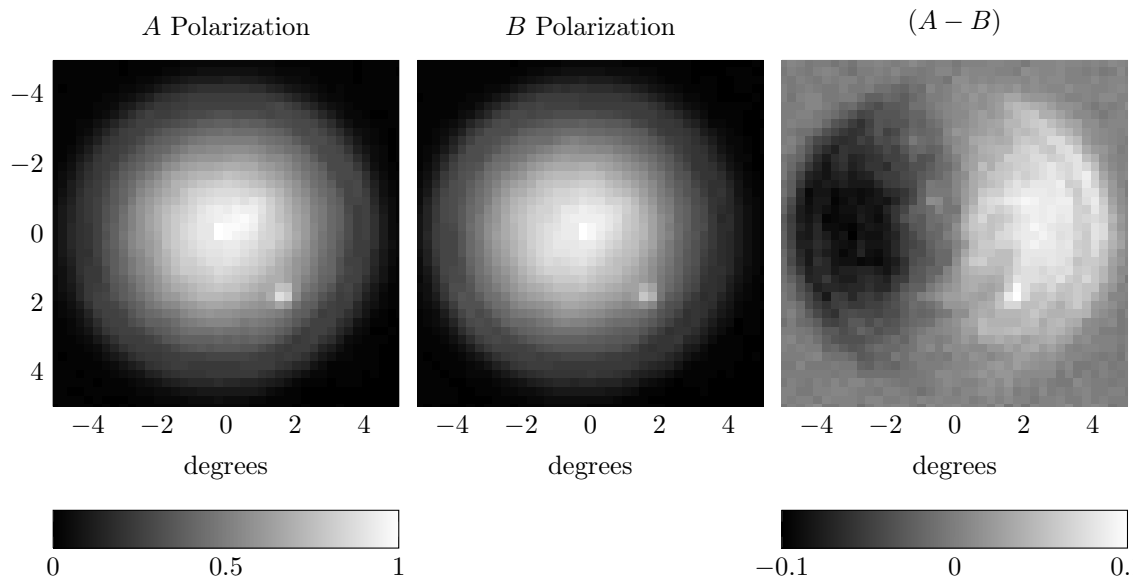


Figure 3.15: Measured beam mismatch in the near field. *Left:* A typical A polarized detector's optical response mapped in the near field, normalized to unit amplitude. *Center:* A typical B polarized detector's near-field response, also peak normalized to unity. *Right:* Difference of A and B. The left-right splitting of the beam centroids was found to be caused by phase gradients in the summing tree. The bright spot in the bottom right quadrant is from secondary reflections within the telescope and is diffusely coupled to the sky.

centroid splitting. Beams measured in the near field show a consistent mismatch in the A and B beam centroids. Moreover, the centroid displacement is consistently (and suggestively) co-aligned with the polarization axes of each tile, and thus also the summing tree axes. The amplitude was measured to be constant across the focal plane, except for small subset of pixels suffering from the severe beam steering illustrated in Figure 3.14. Mismatch in the near-field centroids will not alone lead to any substantial far-field beam mismatch. While the beams may be displaced in the near field, the resulting angular displacement on the sky is negligible. It is true, however, that mechanisms that give rise to phase mismatch in the microstrip summing tree will also tend to give rise to amplitude mismatch, which translates directly to beam mismatch in the far field. Additionally, non-idealities in the optics of the instrument, such as spatially varying birefringence or an out-of-focus system, can complicate the situation, and phase mismatch in the object plane can result in mismatch in the far-field beam performance. In the case of BICEP2, it has not been demonstrated that this near-field mismatch is correlated or directly related to the far-field centroid mismatch described in Section 3.1.4.

Subsequent detector development efforts have reduced the near-field mismatch by adding a phase lag to the summing tree. The additional path length corrects the phase gradient difference between A and B detectors that results from interference within the summing tree. The efforts to improve

the matching of phased-array antenna beams in the aperture plane is described in detail in O'Brient et al. 2012.

3.4 Polarization angle calibration

A key advantage of the BICEP2 experimental approach is the ability to characterize the polarization angles of each detector to high precision using ground-based calibrators. This is made possible by the relatively short far-field range, which is roughly 100 m. Precise polarization angle calibration is critical for constraining potential systematics. Systematic uncertainty in the polarization orientation of the detectors with respect to the sky leads directly to E-to-B mixing, resulting in false EB correlation. As E and T are correlated, this also results in false TB correlation. While a global uncertainty in the effective axes of sensitivity of the instrument is a potentially serious source of systematics, uncertainty in the per-detector polarization angles can likewise be pernicious.

It is important to emphasize that the requirements for calibrating the polarization angles are very different for the BB spectrum than the EB and TB spectra. If we let $\Delta\psi$ be the difference in the assumed polarization angle from the true value, then the leakage in the BB and EE spectra will go as $\sin^2(2\Delta\psi)$. In contrast, the leakage into EB will go as $\sin(2\Delta\psi)$. As a result, the desire to maintain a systematics-free EB spectrum requires much higher calibration accuracy.

As the EB spectrum from simple inflationary models is predicted to be zero, it is possible to tune the effective polarization angles until the EB spectrum is minimized. This is undesirable, however, as other proposed cosmological mechanisms, such as cosmic birefringence (Carroll 1998), can generate a non-zero EB spectrum.

The BICEP2 telescope uses three different methods for determining polarization angles, each susceptible to different systematic uncertainties. The first method uses a thin rotating dielectric sheet placed directly above the window. The second method acquires rasters of a 100% polarized source in the far field at a variety of DK angles. The third method likewise acquires maps in the far field, but with a rotating polarized source. The analysis of the latter two methods are ongoing and, at this stage, less accurate and repeatable than the dielectric sheet measurement. For this reason, our attention will focus on the first of these methods.

3.4.1 Dielectric sheet calibrator measurements

Polarization angles were measured with a beam-filling thin dielectric sheet (colloquially called the "yukical"), rotated about the boresight to produce a polarized signal. An identical technique was used for POLAR (O'Dell 2002) and BICEP1 (Takahashi et al. 2010). The thin film is mounted at 45 degrees relative to the optical axis. The film acts as a polarized beam splitter, reflecting some of the beam into the warm absorptive lining around the splitter and transmitting the rest to the sky.

By rotating the beam splitter relative to the focal plane, the tilt axes of the calibrator rotate with respect to the polarization axes of the detectors, thus creating a periodic polarized signal. As the hardware for the measurement was inherited from BICEP1, we refer the reader to Yuki Takahashi's thesis (2010) for a complete description of the calibrator, appropriately deemed the "Yuki-cal."

The procedure for acquiring polarization angle measurements using the dielectric sheet calibrator is as follows: the apparatus is installed in place of the forebaffle directly above the window of the telescope. Because a substantial fraction of the beam is transmitted to the sky, data are acquired in only the best of weather. A film thickness and index is chosen to provide the requisite signal-to-noise while avoiding gain instability. The detectors are then biased onto either the Ti or Al transitions.[4] Before acquiring the calibration scan, the telescope is "dipped" in elevation to provide an unpolarized signal modulation, from which a relative gain correction between A and B is derived. Scans are acquired by counter-rotating in DK and AZ. The counter-rotation fixes the beam location on the sky, while the calibrator (attached to AZ axis but not the DK axis) rotates about the boresight.

Extracting polarization angles from this measurement requires a full model of the projected axes of polarization sensitivity onto the dielectric sheet. This model is presented in detail in Appendix B. This is based on a similar model presented in Takahashi 2010. Unlike the Takahashi model, we assume at the outset that the deviation of the polarization axes of the detector pair from nominal can be represented by a single angle α, which is equivalent to assuming perfect orthogonality between the detectors. This is a reasonable approximation: Measurements of the cross-polar beam in the far field constrain cross-polar response to < 0.005. If we assume a simple sinusoidal model, this implies that the axes must be co-aligned to < 0.28 degrees. This simplifying assumption allows us to fit a single polarization angle to the pair-difference signal. By assuming orthogonality, we can reject the common-mode atmospheric fluctuations and enormously improve the polarization angle fits. Fitting the response of each detector separately requires stability in the common-mode signal beyond what we were able to achieve during the austral summers.

The model relies on a few externally measured quantities: the tilt of the dielectric sheet, the sheet material properties (including the index of refraction and thickness), and the ray angles of each of the detectors. The tilt of the sheet was measured with respect to gravity with a digital level before and after each scan (represented by t in the appendix, and close to 45 degrees). The agreement between the beginning and end of the scan was typically < 0.03 degrees. The lateral tilt across the surface of the dielectric sheet (represented by γ in the appendix) was also measured with a digital level meter. Between installations of the calibrator, the value of γ was observed to change by as much as 1.5 degrees, but during a measurement was repeatably measured to < 0.02 degrees. The sheet thickness was measured in the lab, while the index of refraction was taken from external

[4] In the end, the data acquired on the Al transition proved less susceptible to gain compression than the Ti data, though both data sets were used in analysis.

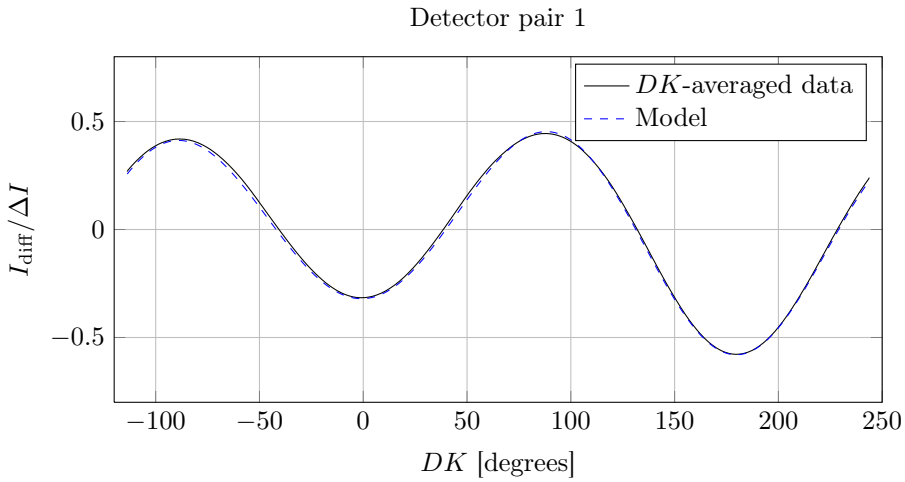


Figure 3.16: Dielectric sheet calibrator pair-difference data and fitted model for Detector Pair 1. The detectors are first relatively calibrated off of the atmosphere by performing a dip in elevation, then differenced to construct the pair-difference irradiance, $I_{\rm diff}$. The pair-difference signal, normalized by the brightness difference between the warm absorber and the sky, ΔI (a free parameter in the model), is extremely well-matched by the model of the polarized signal from the dielectric sheet, the details of which can be found in Appendix B.

sources. The fit angle α is a weak function of both the index and the sheet thickness, and thus does not present significant uncertainties. The ray angles of the detectors are taken from the detector centroid fits in the far field, as described in Section 3.1.3. With these external inputs accurately measured, the model leaves only two free parameters. The first is α, which is the angle of the A and B polarization axes relative to DK. The second free parameter is ΔI, the amplitude of the signal, which is proportional to the difference in temperature between the absorptive lining and the sky temperature at zenith.

Polarization angles are derived from a total of 5 independent measurements separated in time. These were acquired in August 2010, November 2010, March 2011, November 2011, and December 2012. The first three measurements used a 2 mil thick mylar film while the final two measurements were taken with a thinner 1 mil sheet. To combine the measurement results, weights were derived from the inverse variance of the residual after subtracting the fitted model, thus down-weighting noisy or poorly modeled data. Using this weighted combination, we find a best-fit global rotation of the polarization angles $\Delta\chi = -0.18$ degrees. This is in excellent agreement with an overall rotation of the focal plane inferred from beam centroids, which was fitted using CMB data to be equal to -0.17 degrees. Among the five scans, the maximum deviation from the mean was $+0.1$ degrees, but this scan also contained the least weight. We therefore quote a conservative estimate of the 1-σ uncertainty in the global rotation of the polarization angles to be $0.1/\sqrt{5} = 0.04$ degrees. This uncertainty is a significant improvement over BICEP1, which achieved a polarization orientation

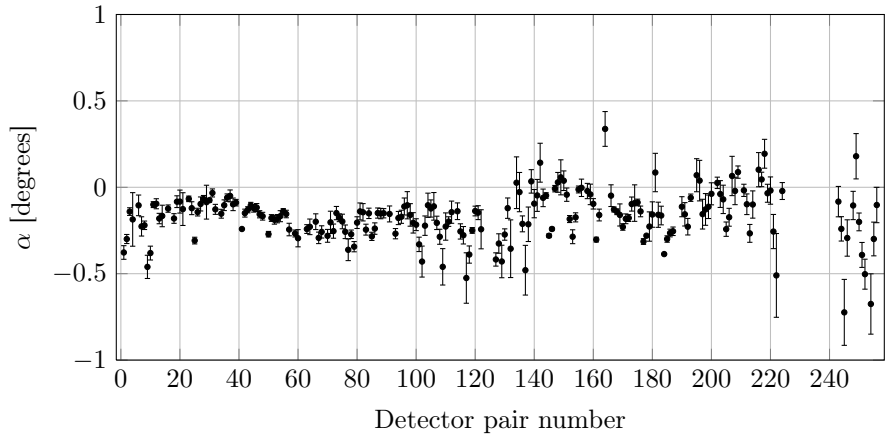


Figure 3.17: BICEP2 measured polarization angle deviation with associated measurement uncertainty. The angle α corresponds to the deviation of the polarization axes from nominal, as defined in Equation B.15. These measurements are consistent with pointing center fits, which measure a global rotation of the focal plane of -0.18 degrees from DK. Error bars are equal to the square root of the weighted variance.

uncertainty of $\Delta\chi < 0.7$ degrees (Takahashi et al. 2010). The dominant uncertainty in BICEP1 was an apparent overall offset of 1.0 degree between calibration measurements taken in 2006 and in 2007-2008 that was never fully understood. BICEP2 observed much tighter consistency over repeated independent measurements, thus achieving a much tighter global constraint.

In addition to measuring the global rotation of the polarization axes, we can similarly assess the per-detector scatter and uncertainty. We repeatably measure scatter in the polarization angles across the focal plane: The standard deviation taken across all operational detector pairs (referred to as "really good lights") is 0.14 degrees. We estimate the median 1-σ per-detector polarization uncertainty to be 0.08 degrees. This is calculated as the square root of the weighted variance, where, as before, the weights are derived from the fit residuals.

3.5 Magnetic response

Extensive laboratory measurements of the magnetic response of the detectors and SQUIDs were made prior to deployment with the goal of measuring the system's magnetic sensitivity and validating the shielding design. The magnetic attenuation was measured using a Helmholtz coil, oriented along the optical axis of the telescope ($\hat{z}$), with 15 turns of copper wire at a 78 cm diameter. The coils were arranged such that their vertical separation was equal to the radius (in a true Helmholtz configuration). Measurements of early magnetic shielding configurations were also taken along the $\hat{\imath}$ and $\hat{\jmath}$ axes. These measurements were made more complicated by the fact that the coil separation was much larger than the coil radius (due to mechanical restrictions), introducing some field non-

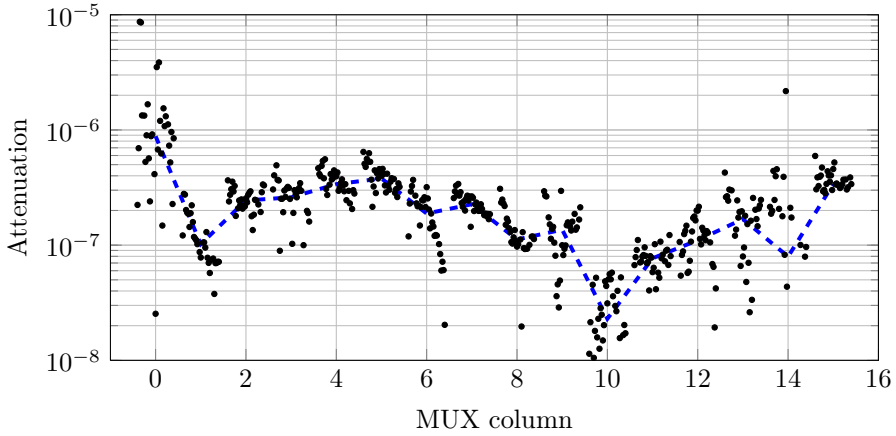


Figure 3.18: Measured end-to-end magnetic field attenuation, measured with the applied field oriented along the optical axis. The attenuation factor assumes an effective area of 882 μm^2, as reported in Stiehl et al. 2011. The magnetic attenuation was measured at a variety of different applied field amplitudes, represented as a cluster of points around each MUX column. The average overall applied field amplitudes is represented by the dashed blue line.

uniformity.

Because of the orientation of the SQUIDs in the focal plane and the orientation of the first-stage magnetic shield, we expect sensitivity principally in the $\hat{z}$ direction. The field strength near the center of a Helmholtz coil is given by:

$$B = \left(\frac{4}{5}\right)^{3/2} \frac{\mu_0 n I}{R}, \tag{3.12}$$

where μ_0 is permeability of free space, n is the number of turns, I is the applied current, and R is the coil radius. Plugging in the values above, we find the applied magnetic field to be 0.34 Gauss per Amp. It is convenient to perform an additional conversion from applied magnetic field to magnetic flux at each successive SQUID stage. This conversion factor is called the effective area, $A_{\rm eff}$, measured to be 882 μm^2 at the first-stage SQUIDs (SQ1) and 483 μm^2 at the second-stage SQUIDs (SQ2) (Stiehl et al. 2011). We can then express the pickup as a fraction of a magnetic flux quantum, Φ_0. For SQ1, we calculate the pickup to be 11.7 Φ_0/A, and 7.8 Φ_0/A at SQ2. We measure the response at the various SQUID stages as current response as a fraction of I_0, the current required to cycle through one Φ_0. As $I_0 = \Phi_0/M$, where M is the SQUID/input coil mutual inductance, a fractional change in the current response is equal to the a fractional change in the magnetic flux (modulo a gain factor).

During the fifth engineering run of BICEP2 at Caltech, we measured the response amplitude at each stage, varying the applied Helmholtz coil current I from 100 mA to 14 A using a Labview-controlled high-current amplifier. By setting the SQUID bias to zero at various stages, it is possible

to differentiate between magnetic pickup at SQ1, SQ2, and the SSAs. In several early engineering runs of BICEP2, we discovered substantial magnetic pickup in the SSAs by setting the SQ1 and SQ2 bias to zero. To contend with this pickup, a multi-layer Metglas wrap was added to the shielding of the SSA modules. This Metglas wrap increased the magnetic attenuation by over an order of magnitude. Similarly, we added a multi-layer Metlas backing behind the SQUID MUX chips.

The achieved end-to-end magnetic attenuation is between -70 to -60 dB (Figure 3.18). This attenuation is somewhat better than the magnetic shielding model predictions (see Section 2.6), but that model did not include the Metglas backing immediately above the MUX chips. Taken together, these measurements are consistent with the Cryoperm shield, Nb spittoon, and Nb backplane, providing a factor of $10^5 - 10^6$ suppression in the $\hat{z}$ component of the magnetic field, with the Metglas backing providing an additional 1 to 2 orders of magnitude.

While the magnetic response of the system was not measured in the field, the magnetic shielding configuration remained virtually unchanged from the system measured at Caltech. We therefore expect similar levels of magnetic attenuation in the as-deployed system.

3.6 Thermal response

The thermal transfer function of the passive filter, a $2.5 \times 2.5 \times 5.5$ cm block of 316 stainless steel between the fridge heat strap and the focal plane, was characterized by pulsing a heater on the heat strap (on the "dirty" side of the filter) and measuring the corresponding temperature response on the focal plane. The filter's characteristic 3 dB point was measured at 0.3 mHz, well outside of our temporal science band (Kaufman et al. 2013a). In this same measurement, the heat capacity was measured as 0.33 J/K. The thermal conductance of the filter (including contact impedance) was separately measured to be 3.0×10^{-4} W/K. Using the measured heat capacity, this yields a characteristic time-constant of $\tau = C/G = 1100$ s, which is consistent with the measured 3 dB point.

The thermal stability of the focal plane is measured using NTD thermometers bonded directly onto the detector tiles. During a CMB scanset (roughly 50 minutes of data), the standard deviation of the temperature on the focal plane is typically 0.3 mK. Full season maps were made with the calibrated NTD thermometry to investigate the level of thermal contamination in the BICEP2 science maps. The thermometry data were processed in the exact same manner as the detector time series, including third-order polynomial filtering and an azimuth-fixed template subtraction. These maps were then processed into power spectra. Through this process, it was determined that thermal fluctuations at $\ell = 100$ are less than 0.4 $\mathrm{nK_{FPU}}$ (at the focal plane), a noise-dominated upper-limit (Kaufman et al. 2013b). The array-averaged thermal responsivity has been measured to be 4.2 $\mathrm{nK_{CMB}/nK_{FPU}}$. We thus find that thermal fluctuations contribute less than 1.6 $\mathrm{nK_{CMB}}$ at $\ell = 100$, nearly a factor of twenty lower than the predicted B-mode amplitude at $\ell = 100$ for

$r = 0.01$.

3.7 Pointing reconstruction

The telescope pointing follows the model presented in both Yoon 2007 and Chiang 2008. We refer the reader to those documents for a complete description. An illustration of various pointing model conventions is provided in Figure 3.19, adapted from Yoon 2007. The "offline" pointing model translates raw encoder counts to the boresight pointing on the sky. The necessary model parameters include an azimuthal and elevation tilt, which are tilts of the telescope axes relative to the local topocentric horizontal coordinates. The pointing model also allows for a flexure term in the azimuthal pointing, but has so far been set to zero. These model parameters are derived from star pointing, acquired with the optical star pointing camera during cryogenic operations (nominally every 3 or 6 days, weather permitting). Celestial objects are tracked by the telescope while a user measures an offset between the nominal boresight center and the celestial target in a digital viewfinder. The model parameters are then fit in offline analysis, given the measured pointing offsets.

The model parameters change gradually over time, which is perhaps not surprising, given that the telescope is housed within a building that sits upon over a mile thick sheet of ice. The fitted tilts suggest a slow sinking of one corner of the building during the winter months that subsequently rebounds during the summer. Despite these kinematics, the star pointing fit residuals are typically 10-12" rms, which is sufficiently small, given the comparatively large beam size.

3.8 Forebaffle loading

To calculate the forebaffle loading, we took a simple data set, consisting of a 20 minute dwell at a fixed azimuth and elevation in good weather, with and without the forebaffle. The difference in the mean value of the detector time series during this period is proportional to the change in loading, due to optical coupling to the warm upper forebaffle. The offsets (in ADU) are converted to fractional airmass by applying a relative calibration derived from el-nods acquired at the beginning and end of the measurement (see Section 4.1.1 for a description of the el-nod calibration).

After applying relative gain corrections, the forebaffle loading in fractional airmass can be converted to K_{CMB} by multiplying by the approximate zenith temperature, 12.3 K_{CMB}. The forebaffle loading is found to be between 4 and 6 K_{CMB}, with significant variation across the focal plane. The median across good light channels is 4.6 K with a standard deviation of 0.73 K (see Figure 3.20). This is significantly higher than the equivalent numbers for BICEP1, which varied between 1 and 2.5 K_{CMB} across the focal plane (with a median of 1.7 K_{CMB}). We attribute this higher loading, at least in part, to the higher illumination of the BICEP2 aperture. For a subset of pixels, the near-field

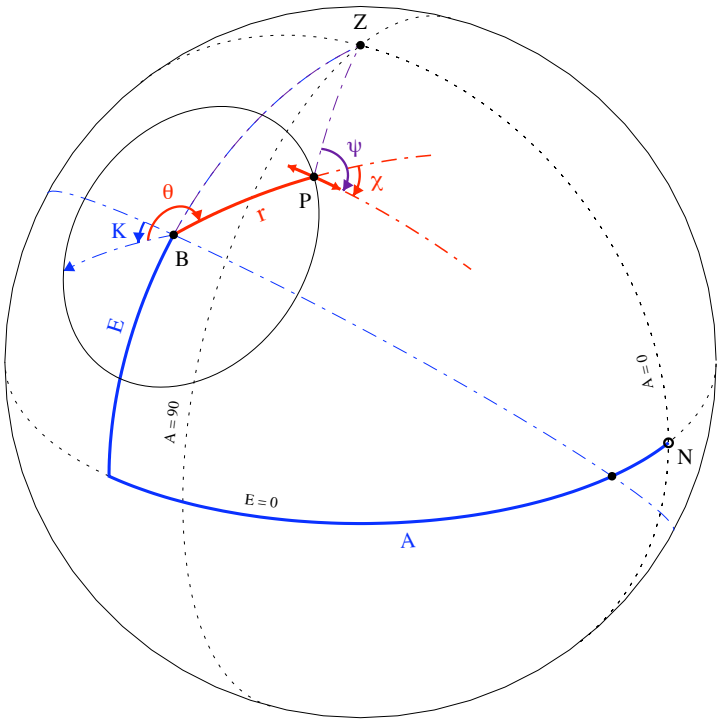


Figure 3.19: Pointing conventions for BICEP1 and BICEP2, adapted from Yoon 2007.

beam steering described in Section 3.3 may also contribute.

3.9 Instrument sensitivity and noise performance

Investigation of the instrument sensitivity and detector noise performance has been reported in detail in Brevik et al. 2010 and Ogburn et al. 2012. Here, we briefly summarize the sensitivity and noise performance of the instrument as reported in these references.

The per-detector noise was assessed by measuring the per-detector time series noise between $0.1 - 1.0$ Hz, which roughly corresponds to the modulation frequency of the ℓ range of interest during normal science observations. Detector noise is measured in noise equivalent current (NEI), which is taken as the square root of the noise power spectral density (PSD) expressed in units of $\mathrm{A}/\sqrt{\mathrm{Hz}}$. This is converted to noise equivalent power (NEP) by simply multiplying by the voltage across the TES, V_{TES}. To convert NEP to noise equivalent temperature (NET), we multiply by the optical responsivity. This is measured using a Rayleigh-Jeans source, and correspondingly expressed in terms of $\mathrm{pW}/\mathrm{K_{RJ}}$. We can convert to $\mathrm{K_{CMB}}$ by a factor of 1.7, which just comes from performing an integral of the blackbody function over a 2.7 K blackbody over our spectral band and dividing by the equivalent integral for a Rayleigh-Jeans source. The measured median per-detector NET for

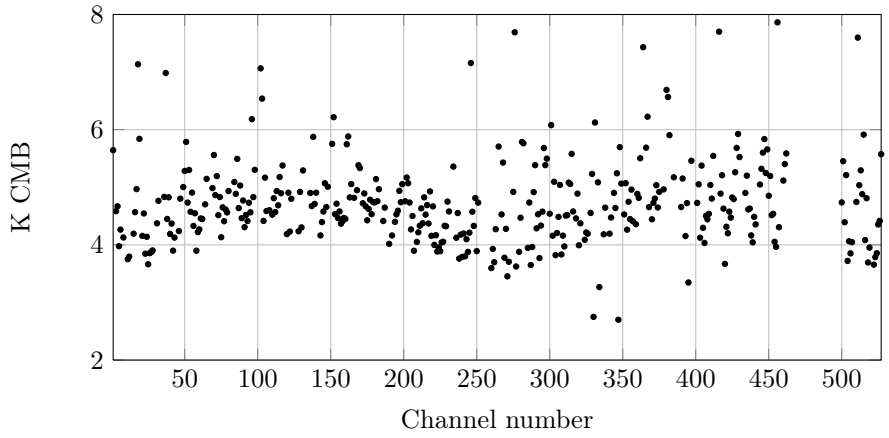


Figure 3.20: Measured forebaffle loading in BICEP2 from optical coupling of the detectors to the warm upper forebaffle. An approximate measurement of the loading is taken with two ∼20 minute integrations with and without the forebaffle, and simply converting the baseline shifts to equivalent units of K_{CMB} using the appropriate relative gain corrections and an assumed zenith temperature of 12.3 K_{CMB}.

BICEP2 during the 2010 season is 433 $\mu K_{CMB}\sqrt{s}$. Toward the end of the 2010 observing season, the per-detector NET was reduced to 379 $\mu K_{CMB}\sqrt{s}$, made possible by optimizing the TES bias (summarized in Brevik et al. 2010). The NET was further reduced for the 2011 and 2012 observing seasons by increasing the readout rate from 15 kHz to 25 kHz. By increasing the readout bandwidth, it was possible to filter the out-of-band aliased noise, the result of which was to reduce the median per-detector NET to 316 $\mu K_{CMB}\sqrt{s}$. This can be converted to an instrument NET by simply scaling by the number of operational detectors, yielding an instrument sensitivity of 15.9 $\mu K_{CMB}\sqrt{s}$ (for the 2011 and 2012 observing seasons).

The instrument sensitivity has also been assessed using scan-direction jackknife maps, whereby left-going and right-going maps are differenced. These maps are constructed for each detector pair separately. These map-based noise estimates yield a median per-detector noise estimate of 422 and 313 $\mu K_{CMB}\sqrt{s}$, for the 2010 and 2011-2012 data selection, respectively (Ogburn et al. 2012). This corresponds to an instrument sensitivity of 21.5 $\mu K_{CMB}\sqrt{s}$ (2010) and 15.8 $\mu K_{CMB}\sqrt{s}$ (2011-2012), in strong agreement with the time series-based noise estimates. Detailed discussions of the various contributions to the detector noise can be found in Brevik et al. 2010, Brevik 2012, Kernasovskiy et al. 2012, Ogburn et al. 2012, and references therein.

Chapter 4

Analysis Methods for BICEP1 and BICEP2

In this chapter, we will summarize the analysis methods used to generate temperature and polarization power spectra from time-ordered data (TODs), which are then used for cosmological parameter estimation. The primary BICEP2 Matlab[1] pipeline is based on a MASTER-style analysis (as described in Hivon et al. 2002). The primary analysis code was taken directly from the BICEP1 Matlab pipeline. This, in turn, was based on the QUAD pipeline, as described in Pryke et al. 2008. As a result, many of the analysis techniques used at various stages in the pipeline, including map-making and power-spectra generation, carried over directly from BICEP1. Other parts of the pipeline underwent development for BICEP2-specific challenges. This includes an expanded cut structure, a more sophisticated instrumental polarization simulator, and new analysis techniques for regressing various sources of instrumental polarization from data.

In the following sections, we concurrently describe the BICEP1 and BICEP2 analysis pipelines, calling particular attention to the ways in which they differ. New developments in both pipelines have led to new scientific results, which we describe in the final chapter of this thesis.

In the first section of this chapter, we give a broad overview of the data reduction procedure, including all of the steps required to generate calibrated temperature and polarization maps. In the second section, we describe the process whereby spectral bandpower estimates are derived from our observed maps. In the third section, we describe the simulation pipeline. This is used for our MASTER-style analysis, including noise de-biasing, bandpower estimation, and parameter estimation. The fourth section will focus on new analysis techniques developed for generating, and then subsequently removing, potential sources of instrumental polarization resulting from gain and beam mismatch. In the fifth section, we review the quadratic estimator construction for r. Finally, the last section will review the methods used to generate likelihood estimators using simulated bandpowers.

[1] `http://www.matlab.com`

4.1 Data reduction

The primary BICEP1 and BICEP2 data reduction can be roughly divided into four steps, each handled by separate parts of the analysis pipeline: low-level TOD reduction, map construction, data selection, and then final co-addition. The pipeline is rich with project-specific jargon, which we will briefly introduce. In both experiments, the data is divided into half-scans (consisting of a single azimuthal sweep across the observing region, or about 1 minute of data), a scanset (consisting of a full set of half-scans at a fixed azimuth, about an hour of data), a phase (consisting of half of the elevation coverage of our observing region at a single DK angle, roughly 10 hours of data), and, finally, a schedule (a full execution of all phases at a single DK angle, roughly 3 days of data). For BICEP1, reduced TODs are produced for every phase, whereas for BICEP2 they are produced for every scanset. These TODs are subject to some data cuts, and are then binned into pairwise maps (called pairmaps). Finally, pairmaps are combined over multiple scansets (phases, for BICEP1) into co-added T, Q, and U maps. It is these co-added maps from which power spectra are generated.

We review these steps in detail below, beginning with the low-level data reduction.

4.1.1 Low-level reduction

The analysis begins with raw data files, which are stored as time-ordered data in ADU units, for our purposes, an arbitrary unit from digitization proportional to the TES current. The first step in the low-level reduction is the deconvolution of the time streams. There are two transfer functions that need to be taken into account: the filter applied by the readout electronics to reduce out-of-band noise and the detector transfer functions. For BICEP2, there are two filters applied by the readout electronics. The first is a causal Butterworth filter applied by the MCE with a nominal 3 dB point of 137 Hz, and the second is an acausal FIR filter applied by the data acquisition computer before downsampling. Because of the extremely fast response time of the BICEP2 detectors, we have thus far found that the detectors' transfer functions can be treated as infinitely fast. Details of the deconvolution can be found in Brevik 2012. For BICEP1, the comparatively slow NTD time constants could not be ignored and had to be measured in detail, an effort that is well-described in Yoon 2007.

After deconvolution, the time streams are "de-skipped." During BICEP2 data taking, samples are occasionally skipped by the data acquisition computer, leaving gaps that are typically single samples in the TOD. These occasional skipped samples are handled by flagging the data and interpolating over the gap.

In the next step of low-level processing, a de-glitching algorithm removes anomalous delta function spikes and discontinuous steps. The de-glitching procedure in BICEP2 is more complex than in BICEP1, largely because of the presence of data pathologies unique to BICEP2's amplification and

multiplexing schemes. The glitches seen by BICEP2 can be caused by a wide variety of phenomena, some of which are understood, while others are not. Known causes of glitches include i) flux jumping, where a SQUID "jumps" from one stable lock point to another, causing a stepwise discontinuity, ii) cross-talk induced by flux jumping, iii) snow falling on the window, causing a common-mode step, iv) cosmic rays, which are suspected to cause flux jumps, and v) modulation of the membrane film above the window. The de-glitching routine handles these various phenomena differently, but generally operates by removing the brief transient and the subsequent ringing from the digital filters. In the case of stepwise discontinuities, the de-glitcher also attempts to re-level the detector time series before and after the step.

Following these steps, the offline pointing model is called to reconstruct the pointing, as described in detail in Yoon 2007 and Chiang 2008. This converts the telescope encoder counts into the boresight pointing on the sky in topocentric horizontal coordinates.

After deconvolving, de-skipping, and de-glitching the scan data, the calibration data are reduced. This begins with flagging the calibration data acquired at the beginning and end of each scanset, along with telescope turnarounds. Next, relative gain corrections are calculated for each detector, in a procedure shared by BICEP1 and BICEP2. The relative gain is measured by "nodding" the telescope in elevation by 1.2 degrees peak-to-peak (called "el-nods"). This produces a small ($\sim$ 100 mK) unpolarized signal in both A and B. The detector time series are then regressed against the secant of the elevation encoder trace to derive a relative gain. Further details on the relative gain correction procedure are available in Appendix A.1. These el-nods are performed every 50 minutes, at the beginning and end of each scanset. The mean of the relative gain corrections derived from the leading and trail el-nods are applied to the detector TODs. These calibrated TODs are then saved to disk.

In addition to the el-nods, BICEP2 also acquires load curves at the beginning and end of each scanset. These load curves are used to measure the operational resistance and the Joule power of the TES, from which the optical loading can be inferred. The load curves are partial, in that they do not sweep through the entire bias range of the TES, but instead stop at the nominal operational TES bias point. The derived calibration data are stored together with the other standard data products. These and a few other quantities (such as the scanset standard deviation) are used for data selection in later parts of the pipeline.

4.1.2 Data selection

Data selection is one area of the pipeline that is highly divergent between BICEP1 and BICEP2. While BICEP1 relied on a single cut criterion to exclude bad data, BICEP2 has expanded the cut framework to include many more statistics with which to evaluate data quality. This is in part due to data pathologies unique to BICEP2, but also by the desire to recover data by using cuts with

finer granularity. The BICEP1 cut is derived from the relative gain standard deviation as a statistic to identify and remove the worst weather. If the median standard deviation of the relative gains exceeded 20% of the average, the entire phase was cut. A full description of the cut as defined and applied in BICEP1 can be found in Chiang 2008.

In BICEP2, data cuts occur at two stages. So-called "Round 1" cuts occur before co-adding TODs into pairmaps (constructed separately for each scanset). As a result, these cuts can have dimensions of $n_{\mathrm{HS}} \times n_{\mathrm{chans}}$, where n_{HS} is the number of half-scans and n_{chans} is the number of channels[2]. "Round 2" cuts, on the other hand, occur at the co-addition stage and thus have dimensions of $1 \times n_{\mathrm{pairs}}$ (meaning that entire scansets or channel pairs can be excluded, but not individual half-scans). The current implementation of the cut scheme for BICEP2 is summarized in Tables 4.1 and 4.2.

[2]We are somewhat casual about distinguishing between cuts that exclude individual channels versus cuts that exclude channel pairs. Since excluding a channel within a pair will cause its partner to be cut as well, this distinction makes no practical difference.

Cut name	Description[1]	Dimensions[2]	Threshold	Unique cut fraction[3]
`fb_nancount`	Number of NaNs	$n_{\rm HS} \times n_{\rm chans}$	1	8×10^{-4}
`fb_std_p0`	Std. dev. P0 filtered TODs	$n_{\rm HS} \times n_{\rm chans}$	Inf	0
`fb_std_p0_darks`	Std. dev. P0 dark TODs	$n_{\rm HS} \times n_{\rm darks}$	5	0
`fb_std_p3`	Std. dev. P3 filtered TODs	$n_{\rm HS} \times n_{\rm chans}$	Inf	0
`fb_std_sd_p0`	Std. dev. unfilt. sum/diff TODs	$n_{\rm HS} \times n_{\rm chans}$	Inf	0
`fb_std_sd_p3`	Std. dev. filt. sum/diff TODs	$n_{\rm HS} \times n_{\rm chans}$	Inf	0
`fb_std_uncal`	Std. dev. filt. sum/diff TODs before cal	$n_{\rm HS} \times n_{\rm chans}$	Inf	0
`is_fj_row`	Is flux jumping? (row)[4]	$n_{\rm HS} \times n_{\rm rows}$	Inf	0
`is_fj_col`	Is flux jumping? (column)[5]	$n_{\rm HS} \times n_{\rm cols}$	Inf	0
`syncsampnum_diff1`	Flag for out-of-order sample	$n_{\rm HS} \times 1$	[logical]	0.021
Cut-on-cuts				
`passfrac_col`	Frac. of channels cut within MUX column	$n_{\rm HS} \times n_{\rm cols}$	0.7	0
`passfrac_nchan`	Frac. of channels cut across FP	$n_{\rm HS} \times 1$	0.7	0

1. P0 and P3 are poly-filter operations of zeroth and third order, respectively.
2. Here $n_{\rm HS}$, $n_{\rm chans}$, $n_{\rm darks}$, $n_{\rm rows}$ and $n_{\rm cols}$ are the number of half-scans, channels, dark channels, MUX rows and MUX columns, respectively.
3. Unique cut fraction is calculated for the entire 2010-2012 data selection.
4. Flag entire MUX row if a detector within that row flux jumps.
5. Flag entire MUX column if a detector within that column flux jumps.

Table 4.1: Summary table of BICEP2 "Round 1" cut definitions, applied before co-adding over detector pairs. Many of the Round 1 statistics exist solely as data diagnostic tools, and are not actually used to down-select data. These statistics are identified as having "Inf" cut thresholds.

Cut parameter	Description[1]	Dimensions[2]	Threshold	Unique cut fraction[3]
`elnod_mean`	el-nod mean	$1 \times n_{\rm chans}$	$2000 - 8000$ ADU	0.0112
`elnod_fracdel`	el-nod change during scanset	$1 \times n_{\rm pairs}$	0.3	0.003
`elnod_abba`	ratio-of-ratio, before/after	$1 \times n_{\rm pairs}$	0.1	0.0001
`elnod_median`	median el-nod responsivity	1×1	$3000 - 6000$ ADU	0.001
`elnod_nancount`	Number of NaN's in el-nod	$1 \times n_{\rm chans}$	1	0.0083
`elnod_gof`	El-nod goodness-of-fit	$1 \times n_{\rm chans}$	Inf	0
`elnod_chisq_diff`	El-nod fractional residual	$1 \times n_{\rm chans}$	Inf	0
`fb_wn_sd_p0`	Mean pair sum/dif PSD from $1.5 - 2.0$ Hz	$1 \times n_{\rm chans}$	Inf	0
`fb_1f_sd_p0`	Mean pair sum/dif PSD from $0.1 - 0.3$ Hz	$1 \times n_{\rm chans}$	Inf	0
`skewness_dif`	Pair-difference skewness	$1 \times n_{\rm pairs}$	0.2	0.0034
`fp_cor`	Median detector-detector correlation (pair-diff)	1×1	Inf	0
`scanset_std`	Std-dev over entire scanset	$1 \times n_{\rm pairs}$	2.5 ADU	0.0086
`stationarity_ab`	Std-dev. of `fp_std_p3` (indiv. det's)	$1 \times n_{\rm pairs}$	0.7	0.0047
`stationarity_dif`	Std-dev. of `fp_std_p3` (pair diff)	$1 \times n_{\rm pairs}$	0.2	0.0041
`tfpu_mean`	Mean focal plane temp	1×1	$0.2 - 0.3$ K	0
`tfpu_std`	St-dev. focal plane temp	1×1	5×10^{-5}	0.0021
`num_fj`	Number of flux-jumps	$1 \times n_{\rm pairs}$	5	0
`num_destep`	Number of glitch-recovered steps	$1 \times n_{\rm pairs}$	5	0.0135
`max_fj_gap`	Max flux-jump step	$1 \times n_{\rm pairs}$	1000 ADU	0.028
`rtes_frac`	TES fractional resistance	$1 \times n_{\rm pairs}$	$0.1 - 0.95$	0.0032
Cut-on-cuts				
`passfrac_halfscan`	Frac. of half-scans passing Round 1	$n_{\rm HS} \times n_{\rm cols}$	0.9	0
`passfrac_scanset`	Frac. of data in a scanset passing Round 2	$n_{\rm HS} \times 1$	0.5	0.0215

1. P0 and P3 are poly-filter operations of zeroth and third order, respectively.
2. Here $n_{\rm chans}$ and $n_{\rm pairs}$ are the number of individual channels and detector pairs respectively. See footnote in Section 4.1.2.
3. Unique cut fraction is calculated for the entire 2010-2012 data selection.

Table 4.2: Summary table of BICEP2 "Round 2" cut definitions, applied to pairwise maps constructed per-scanset before co-addition into maps. As in the case of the "Round 1" cuts, not every cut statistic is used for data selection. These are indicated as having "Inf" thresholds.

4.1.3 Map construction

In the next step in the data analysis, reduced TODs are gathered and binned into pairwise maps, called "pairmaps." The map-making step itself encompasses several other simultaneous analysis steps, which we summarize here. To begin, the cut parameters summarized in Tables 4.1 and 4.2 are calculated from the TODs. Immediately after, "Round 1" cuts are applied, and "Round 2" cut statistics are stored for later use. Next, the TODs (already relative-gain calibrated) are sum and differenced.

At this stage, the pipeline allows for flexibility in choosing the filtering that is applied to the TODs. The default option for BICEP1 and BICEP2 is to first apply a third-order polynomial filter and then apply "ground subtraction." The ground subtraction process begins by accumulating the pair-wise sum and difference TODs into azimuthal bins (which are fixed with respect to the ground). This serves as a ground-fixed template, which is then regressed against the data, thereby removing any signal fixed with respect to the ground. Sky-fixed modes that are comparable to the azimuthal scan width are also heavily filtered.

After filtering, data weights are calculated prior to map binning. Data weights are calculated as the inverse variance of the pair-sum and difference time series. Here too the pipeline offers some degree of flexibility, with the ability to calculate weights over an entire scanset or over individual half-scans (the default is over a scanset).

The next step is the map construction itself. Individual pairwise maps are accumulated for each scanset (consisting of roughly 10 hours of data). Maps are accumulated separately for positive- and negative-going slews. These are used to construct scan-direction jackknives at a later stage. Maps are binned into 0.25×0.25 degree pixels, constructed in Right Ascension/Declination (RA/DEC).

The accumulated quantities associated with the pair-sum data include the weighted pair-sum data itself, the applied weights, the data variance, and various quantities used to track the total integration time. The accumulated pair-difference quantities are more complicated, since they must encode the polarization angle orientation for later conversion into Q and U.

Because the polarization of the CMB is characterized by three numbers (I, Q, and U), and any given observation only provides two numbers (the signal in A and B), multiple DK observations are required to measure both Q and U. For this reason, it is not possible to accumulate pairwise maps into Q and U directly. Instead, the pair-difference data, with applied sinusoidal corrections that depend on the A and B polarization angles, are accumulated and saved to disk.

Descriptions of the inversion of pair-difference data to recover Q and U are available from a large number of sources, including Chiang et al. 2010, Jones et al. 2007, Pryke et al. 2008, and references therein. Here we provide only a brief review of the construction of Q and U from pair-difference data, following closely to the description presented in Jones et al. 2007. We calculate $d_{f,j}$, the pair-difference TOD for the j-th time sample (for a given detector pair) from the uncalibrated TODs of

A and B (denoted as $d_{A,j}$ and $d_{B,j}$) after applying relative gain corrections (η_A and η_B):

$$d_{f,j} = \frac{1}{2}\left(\frac{d_{A,j}}{\eta_A} - \frac{d_{B,j}}{\eta_B}\right). \tag{4.1}$$

The relative gain correction procedure is described in more detail in Appendix A.1. We also define the following angular coefficients:

$$\alpha_j = \gamma_A \cos 2\psi_{A,j} - \gamma_A \cos 2\psi_{B,j} \tag{4.2}$$

$$\beta_j = \gamma_A \sin 2\psi_{A,j} - \gamma_B \sin 2\psi_{B,j}. \tag{4.3}$$

Here ψ is the time- and DK-dependent polarization orientation angle projected onto the sky, as illustrated in Figure 3.19, and γ is the usual polarization efficiency correction:

$$\gamma \equiv \frac{1-\epsilon}{1+\epsilon}. \tag{4.4}$$

As before, ϵ is the "polarization efficiency." It can be thought of as the ratio between the minimum and peak response to a linearly polarized source. An ideal detector will have $\epsilon = 0$. If γ is the

The pair-difference data are related to Q and U through the following matrix equation:

$$\begin{pmatrix} d_j\alpha_j \\ d_j\beta_j \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \alpha_j^2 & \alpha_j\beta_j \\ \alpha_j\beta_j & \beta_j^2 \end{pmatrix}\begin{pmatrix} Q \\ U \end{pmatrix}. \tag{4.5}$$

We introduce the shorthand:

$$\mathcal{P}_j = \begin{pmatrix} \alpha_j^2 & \alpha_j\beta_j \\ \alpha_j\beta_j & \beta_j^2 \end{pmatrix}. \tag{4.6}$$

Note that for a single DK angle, the determinant of $\mathcal{P} = 0$, so $\mathcal{P}$ is not invertible. This matches our earlier claim: To fully measure both Q and U, multiple DK angles are required. For this reason, we save the equivalent of $d_j\alpha_j$ and $d_j\beta_j$ at the pairmap level (acquired at a single DK angle) and invert after co-adding over multiple DK angles.

4.1.4 Map co-addition

The map co-addition stage begins with loading pairmaps (accumulated over scansets) and applying the "Round 2" cuts, summarized in Table 4.2. (This is not the case for BICEP1, which has only a single cut). At this stage, the channel selection is applied, and channels with known sources of polarization contamination, elevated noise, or unstable lock points are discarded. In addition to the usual data accumulation, maps are also acquired in a variety of data "jack-knives", where the data is split by scan direction, time, DK angle, and EL coverage, as well as a variety of channel selection

criteria. The jack-knife maps are acquired separately for each half and then differenced at a later stage in the analysis. After the maps have been accumulated, Q and U maps are calculated by effectively inverting Equation 4.5. The pair-sum data are accumulated into T (temperature maps) as well.

4.1.5 Absolute calibration

After co-addition, the resulting maps are in units of ADU (analog-to-digital unit). To convert these maps to $\mathrm{K_{CMB}}$, we apply the same absolute calibration technique described in Chiang et al. 2010, which we briefly summarize here.

Absolute calibration is derived from cross-correlating observed maps with "re-observed" WMAP maps. To begin, `Healpix` WMAP maps are beam corrected and then re-smoothed to the BICEP beam resolution. The `Healpix` maps are then "re-observed" by processing the maps through the BICEP simulation pipeline. Because the maps rely on an accurate beam correction, there is some degeneracy between error on the assumed beam size used for re-smoothing and the absolute calibration factor, described in detail in Chiang et al. 2010. This error is taken into account in later stages of the likelihood analysis.

The cross-correlation is calculated in multipole space (as in Chiang et al.) as:

$$\tilde{g}^b_{\mathrm{abs}} = \frac{\Sigma_\ell P^b_\ell \langle a^{\mathrm{WMAP1}}_{\ell m} a^{*\mathrm{BICEP}}_{\ell m} \rangle}{\Sigma_\ell P^b_\ell \langle a^{\mathrm{WMAP1}}_{\ell m} a^{*\mathrm{WMAP2}}_{\ell m} \rangle}. \tag{4.7}$$

Here, P^b_ℓ is a top-hat binning operator and WMAP1 and WMAP2 are two different WMAP frequencies. For Chiang et al. the Q-band (40 GHz) and V-band (70 GHz) maps were used. The gain calibration is calculated for each bin b. Over the multipole range of $21 \leq \ell \leq 335$, the gain calibration is approximately flat. A single absolute calibration factor is calculated by taking an average across bins over a limited range of ℓ, restricted to multipoles of low statistical uncertainty. In the case of Chiang et al. this was in the range of $56 \leq \ell \leq 265$. After taking this average over bins, we arrive at $\tilde{g}_{\mathrm{abs}}$ the absolute calibration, which has units of ADU/μK. We denote the absolute calibration with a tilde because the absolute calibration is calculated for maps that have been the el-nod corrected (see Appendix A.1 for additional details). The denominator requires two maps with independent noise so as to avoid a noise bias in the normalization.

4.2 Bandpower estimation

Bandpower estimation proceeds from calibrated maps in several steps: First, E and B raw power spectra are calculated from the two-dimensional fast Fourier transform (FFT) of Q and U (denoted $\tilde{Q}$ and $\tilde{U}$) and collapsed into 1-D angular power spectra as a function of multipole ℓ. Next, the power

spectra are de-biased, for both noise and for E-to-B leakage from the pipeline's imperfect E/B separation. Finally, bandpower window functions and suppression factor corrections are calculated for each of the spectra. We step through this procedure in detail.

4.2.1 Power spectra calculation

With the T, Q, and U maps in-hand, we proceed with angular power spectra generation. The maps are masked with a variance map, which is constructed from the time series variance. Next, two-dimensional FFTs of Q and U are calculated from co-added maps. In this step, a flat-sky approximation is made. We treat the RA/DEC-binned map as a set of rectilinear coordinates, taking the Fourier transform in the x, y plane. The choice to use a flat sky approximation is driven largely by convenience. Power spectrum estimation in the flat-sky limit is extremely simple, and does not require the complicated mathematics of spherical geometry. The flat sky approximation comes at minimal cost, since the resulting E/B mixing can be subtracted as a simple bias with additional variance that is small compared to the uncertainty due to noise.

The E-mode and B-mode map transforms are calculated in the Fourier domain. Letting χ equal the angular position in the u, v plane (measured from North-going modes towards East-going modes, as per the International Astronomical Union convention), we calculate the Fourier transform of the E and B maps as:

$$\tilde{E} = \tilde{Q}\cos(2\chi) + \tilde{U}\sin(2\chi) \tag{4.8}$$

$$\tilde{B} = \tilde{Q}\sin(2\chi) - \tilde{U}\cos(2\chi). \tag{4.9}$$

A heuristic description of the origin of these expressions can be found in Section 1.5.3.

After calculating $\tilde{T}, \tilde{E}$ and $\tilde{B}$, we collapse the 2-D FFTs into 1-D auto- and cross-spectra. Auto-spectra are generated by multiplying the 2-D FFT by its complex conjugate (calculated as $XX^\dagger$, where X may be T, E, or B) and then summing modes over annular bins centered at $(0, 0)$ in the u, v plane. This is effectively collapsing over m, so the resulting function depends only on the multipole moment, ℓ. Cross-spectra are calculated the same way, summing modes in annular bins of $XY^\dagger$. Here X and Y represent possible combinations of T, E, and B. Also, as per the usual CMB conventions, we multiply by a factor of $\ell(\ell+1)/2\pi$ to arrive at the uncalibrated power spectra:

$$\mathcal{C}_b^{XY} = \frac{\ell(\ell+1)}{2\pi} C_b^{XY}. \tag{4.10}$$

The b subscript indicates that these are binned power spectra (to be distinguished from $\mathcal{C}_\ell$, which denotes the underlying theory spectrum).

This procedure for calculating the E and B power spectra results in imperfect E/B separation

for two reasons. First, the finite sky coverage results in discretely sampled bins in the u, v plane. Because of this pixelization, modes that would otherwise be pure E "bleed" into B, and vice-versa[3]. Second, map distortion from the flat sky approximation produces some low-level mixing of E and B. As mentioned previously, this can be accounted for as a simple bias with tolerable additional sample variance. There are algorithmic solutions to improve E/B separation in pseudo-$\mathcal{C}_\ell$ estimators; one such solution is described in Smith 2006 and may be included in future analyses of BICEP2 data.

4.2.2 Noise de-bias

Auto-spectra naturally suffer from a noise bias. In order to present meaningful bandpower estimates, we must account for this bias. The correction of the bias is obtained from the simulation pipeline, which is described in detail in Section 4.3. We generate many realizations of simulated noise, which share the same statistical properties as the noise in the real data. We take the average over all realizations to be the noise bias, denoted as $N_{\ell,b}$ (calculated separately for TT, TE, etc.). This de-biasing requires a highly accurate noise model; over- or under-estimating the noise bias will lead directly to systematic uncertainty on bandpower estimates.

4.2.3 E-to-B de-biasing

As described in Section 4.2, the limited sky coverage and the flat-sky approximation invoked by the power spectrum estimator lead to imperfect E-to-B separation. This can be described as an E-to-B "bias" in the sense that signal-only simulations contain non-zero $\mathcal{C}_b^{BB}$ power for zero input BB power. This is not, however, a systematic bias on the BB bandpowers, since it can be precisely accounted for with simulations. This is to be distinguished from sources of systematic bias on BB from E-to-B mixing from instrument systematics, such as polarization angle uncertainty.

To account for imperfect E/B separation, we generate a suite of simulations containing cosmological E-mode power, but no B-mode power (called "E-no-B" sims). Each simulation is an independent realization of the same underlying theory spectrum, $\mathcal{C}_\ell$. (The generation of these sims is described in more detail in Section 4.3). From the noiseless maps, we generate simulated TODs that are processed into maps and power spectra using the same analysis steps that are used for the real data. With the raw BB bandpowers in hand, the mean over all realizations is taken as the E-to-B bias. This is subsequently subtracted from the real data, thereby removing the E-to-B mixing from the imperfect E/B separation. In principle, this should be performed in the opposite direction as well: There is non-zero B-to-E leakage from the power spectrum estimation technique as well. However, because the cosmological BB power spectrum is known to be of much lower amplitude than the EE spectrum, we ignore this effect.

[3]One can imagine the extreme case where the Fourier plane has only four bins. In this case, any power in $\tilde{U}$ can only produce E power.

This procedure does not account for sources of E-to-B leakage, resulting from instrument calibration uncertainty. Similarly, this procedure does not force the observed EB spectrum to zero. While standard cosmological models predict zero E/B correlation, there are more "exotic" cosmological models, such as cosmic birefringence, which predict a non-zero EB power spectrum. Since the input simulations used in the E-to-B de-biasing procedure assume zero B-mode power, any observed EB correlation in the real data will be preserved.

Imperfect E/B separation also results in increased statistical uncertainty on the final BB bandpowers. Since there is statistical uncertainty on EE due to cosmic variance, E-modes that leak into the BB spectrum will contribute statistical uncertainty to the BB estimates. For BICEP1, the additional statistical uncertainty from the imperfect E/B separation is a small fraction of the uncertainty due to noise. However, with BICEP2's additional sensitivity, improved E and B estimation algorithms are motivated by the non-negligible statistical uncertainty in BB resulting from the BICEP1-style E/B estimator. These are currently being explored for BICEP2.

4.2.4 Suppression factor and bandpower window functions

To calculate final, calibrated bandpowers, we must account for filtering and the bandpower window functions. Filtering, from both the data processing and the beam, is taken into account as a single suppression factor at each ℓ bin, calculated as the ratio of the mean of the simulated bandpowers to the input model bandpowers. The bandpower window function describes how power at different angular scales contributes to each ℓ bin b. We use a similar formal definition of the bandpower window function as in Knox 1999[4]:

$$\langle \mathcal{C}_b \rangle = \Sigma_\ell \mathcal{W}_{\ell,b} \mathcal{C}_\ell. \tag{4.11}$$

Here, $\langle \mathcal{C}_b \rangle$ is the expectation value for the observed bandpowers, $\mathcal{W}_{\ell,b}$ is the bandpower window function, and $\mathcal{C}_\ell$ is the input theory spectrum for some fiducial model. We briefly describe the procedure to calculate both the bandpower window function and the suppression factor.

The bandpower window function calculation begins with the mask window function. The mask window function accounts for correlation between ℓ bins due to the limited sky coverage. The mask window function is calculated as in Pryke et al. 2008. The Fourier transform of the map mask is convolved with an annulus in the 2-D Fourier domain, where the annulus is defined by the chosen bin width used for the final power spectra calculation.

The mask window function $M_{\ell,b}$, which is computed strictly from the mask and the bin definitions, does not depend on the choice of filtering of the TODs. Because the filtering varies across a bin, not all modes contribute equally within the bin (as is assumed during the mask window function

[4]The thing we call the bandpower window function, $\mathcal{W}_{\ell,b}$, is actually different from the definition of the bandpower window function $W_{\ell,b}$ defined by Knox 1999. The two are related as $\mathcal{W}_{\ell,b} = W_{\ell,b}/\ell$.

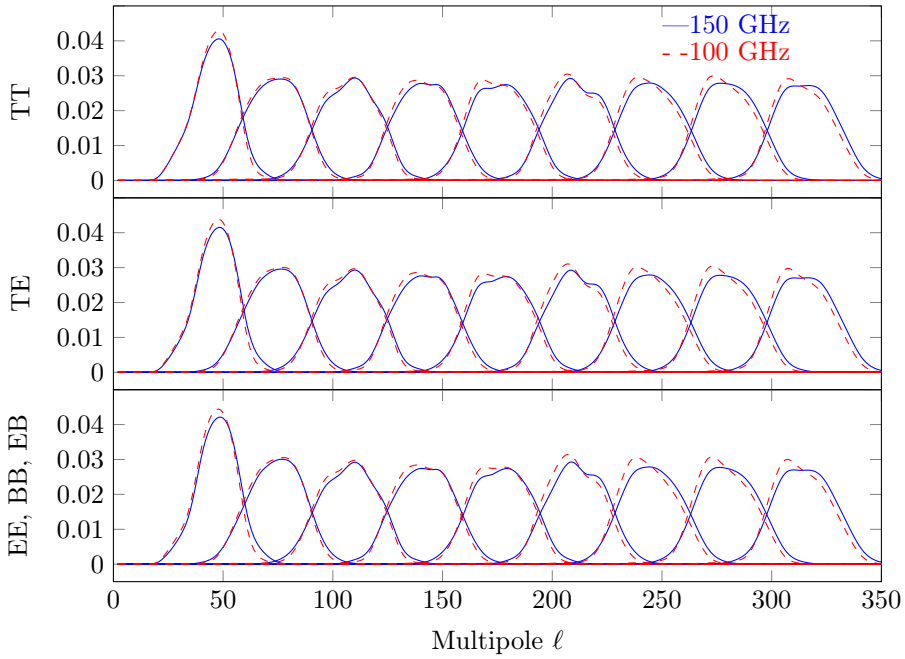


Figure 4.1: Bandpower window functions for the 9 bins used in science analysis in the BICEP1 three-year analysis. The window function for each bin has been normalized to unity.

calculation). There is thus an additional step to compute the "true" bandpower window function from the mask window function. This calculation is performed using the following iterative procedure. To begin, the mask window function is computed to high resolution, typically $\Delta\ell = 1$. Next, a "naive" suppression factor S_b^0 is calculated for each individual bin b as:

$$S_b^0 = \frac{\langle \mathcal{C}_{\ell,b}^{\rm sim} \rangle}{\Sigma_\ell \mathcal{C}_\ell M_{\ell,b}} \tag{4.12}$$

(the '0' superscript will be incremented with subsequent iterations to our estimate of the suppression factor). The numerator is the binned signal-only simulated bandpowers $\mathcal{C}_{\ell,b}^{\rm sim}$ averaged over all realizations for bin b. The denominator is the sum of the input model spectrum (denoted as $\mathcal{C}_\ell$) multiplied by the mask window function $M_{\ell,b}$ for bin b, summed over all ℓ. We call the denominator the "expected value." (Equation 4.12 assumes that the mask window function has been normalized such that $\Sigma_\ell M_{\ell,b} = 1$).

The "naive" suppression factor, calculated at each bin b, is then interpolated via a piecewise cubic Hermite polynomial interpolation. The mask window function is multiplied by the interpolated suppression factor (denoted S_ℓ^0) and re-normalized. In this way, we account for the fact that the filtering is changing across each bin. We then iterate the procedure and recalculate the suppression

factor. For the j-th iteration, the suppression factor is calculated as:

$$S_b^{j+1} = \frac{\langle \mathcal{C}_{\ell,b}^{\rm sim} \rangle \Sigma_\ell S_\ell^j}{\Sigma_\ell \mathcal{C}_\ell M_{\ell,b} S_\ell^j}. \tag{4.13}$$

The denominator can be regarded as the iterated expectation value, that is, the modified expectation value that takes the changing suppression factor into account. The summation $\Sigma_\ell S_\ell^j$ in the numerator is required to not "double count" the suppression of bandpowers. The procedure is iterated until the suppression factor is stationary with subsequent iterations. In practice, only one or two iterations are required; however, we will denote the final suppression factor as S_b^n.

Using the final iterated suppression factor, we can calculate the "true" bandpower window function as:

$$\mathcal{W}_{\ell,b} = \frac{M_{\ell,b} \Sigma_\ell S_\ell^n}{\Sigma_\ell S_\ell^n}. \tag{4.14}$$

Again, we assume $\Sigma_\ell M_{\ell,b} = 1$. Normalizing by $\Sigma_\ell S_\ell^n$ ensures that the true bandpower window function is power preserving. The suppression factor for each bin is then calculated as:

$$S_b = \frac{\langle \mathcal{C}_{\ell,b}^{\rm sim} \rangle}{\Sigma_\ell \mathcal{C}_\ell \mathcal{W}_{\ell,b}}. \tag{4.15}$$

At low ℓ, the suppression factor is dominated by the polynomial filtering and ground subtraction described in Section 4.1.3. At high ℓ, the suppression factor is dominated by the beam roll-off.

As a practical twist on this procedure, some care must be taken in choosing the appropriate bin width. The suppression factor interpolation procedure is made more precise by reducing the bin width (thereby increasing the number of interpolation points). However, this is unattractive for reporting final bandpowers because adjacent bins will be highly correlated. As a result, we use finely spaced bins to calculate the "true" bandpower window function, and afterwards merge the simulated bandpowers, real bandpowers, and bandpower window functions to more coarsely sampled, minimally correlated final bins.

4.2.5 Frequency combination

In the case of BICEP1, there are observations at both 100 and 150 GHz. In order to achieve the highest possible signal-to-noise measurement, bandpowers can be combined across frequencies. In the case of the TT, EE, and BB spectra, there are three unique spectra ($100 \times 100, 150 \times 150$, and 100×150). For the cross-spectra, there is an additional unique spectrum calculated as 150×100. For each ℓ bin, a covariance matrix is calculated from signal-plus-noise simulations across the three (or four) unique frequency combinations. Weights are then calculated as the column-wise sum of

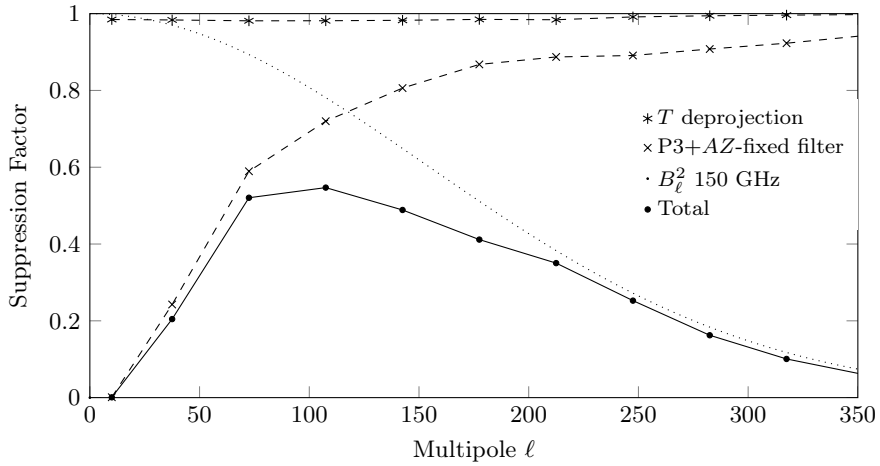


Figure 4.2: Bandpower filter function for EE at 150 GHz, calculated from the suppression factor for the BICEP1 three-year analysis. Low-ℓ attenuation is dominated by the P3 polynomial filtering and AZ-fixed filter, whereas the high-ℓ modes are dominated by the beam roll-off (represented by the square of the beam window function, B_ℓ^2). The T deprojection represents the bandpower suppression resulting from the instrumental polarization regression analysis described in Section 4.4. The TT and BB filter functions are very similar. The first bin, near $\ell = 10$, is not used for science analysis.

the inverse covariance matrix. The weights are then re-normalized such that the sum across all frequencies within a single ℓ bin is unity. These weights are then used to calculate weighted averages for each ℓ bin.

4.3 Simulation pipeline

The simulation pipeline used for both BICEP2 and the BICEP1 three-year analysis are based on the simulation pipeline presented in Pryke et al. 2008 and Takahashi et al. 2010. The simulations serve to: i) account for the noise bias in the auto-spectra, ii) calculate the suppression of bandpowers due to filtering, iii) calculate the E-to-B leakage that results from imperfect E/B separation, iv) report measurement uncertainties on bandpowers, and v) calculate likelihood estimators for r, the tensor-to-scalar ratio. This work was preceded by (and heavily relies on) the MASTER analysis presented in Hivon et al. 2002. In this section, we will describe the mechanics of the simulation pipeline.

4.3.1 Simulated signal generation

Signal-only simulations begin with `synfast` maps generated from `CAMB`[5] input files of best-fit cosmological parameters. For BICEP1, cosmological parameters were extracted from the WMAP five-year analysis (Dunkley et al. 2009). (Future BICEP2 analyses will be able to take advantage of more

[5] `http://camb.info/`

recent analyses from WMAP and Planck). A large number of realizations of T, Q, and U are generated, each with a unique distribution of $a_{\ell m}$'s (499 for BICEP1). The `synfast` maps are smoothed to the focal plane median beam width. Additionally, the first and second spatial derivatives are calculated using the `synfast` package. For BICEP1, all of the maps are calculated at a resolution of `nside` = 512. Future BICEP2 analyses will likely take advantage of higher resolution maps.

Simulated TODs are generated by interpolating the `synfast` maps using the pointing data from the telescope. At each moment in time, each detector's pointing center on the sky is calculated as a reckoning along a bearing angle θ and an angular distance r from the boresight, as defined in Figure 3.19. (The boresight pointing reconstruction is summarized in Section 3.7). Next, the algorithm identifies the `Healpix` pixel center nearest to the detector's pointing center. The value of the map at the true pointing center is then approximated using a Taylor expansion from the nearest `Healpix` center using the first and second spatial derivative maps. This interpolation procedure is used for calculating T, Q and U all in the same manner. Simulations can span the entire observing period of the telescope, or just some subset of data.

The simulated TODs for A and B are assumed to be perfectly matched in gain (unless we explicitly inject mismatched gains), thereby obviating the el-nod relative gain correction step. The TODs, once generated, are carried forward in the analysis in the exact same way as the real data. This includes using the same filtering and weighting procedure. Similarly, the same cuts are applied to the simulated data as the real data. Co-added maps and power spectra are calculated as in Sections 4.1.3 and 4.2, respectively.

Simulated maps containing E-mode power but no B-mode power (E-no-B sims) are used to calculate the E-to-B leakage from the power spectrum estimator (as in Section 4.2.3) and the suppression of EE bandpowers (as in Section 4.2.4). Similarly, B-no-E sims are used to calculate the suppression of BB bandpowers. This simulation set is also used in later stages of the analysis to calculate likelihood estimators for r (described in Section 4.6).

4.3.2 Simulated noise generation

The goal of the noise simulator is to produce TODs that, in both pair-sum and pair-difference, have the same statistical properties as the real data. The noise model is based on the simulator reported in Pryke et al. 2008, but has been somewhat modified for the BICEP2 and BICEP1 analyses.

The noise model is constructed in two different steps: The first step measures and then simulates the low-frequency polynomial modes, while the second step measures and subsequently simulates high-frequency Fourier modes. To begin, all scans within a scanset are concatenated and filtered with a simple first-order polynomial filter. From the polynomial coefficients (the mean and the slope), a joint covariance is calculated between all channels and modes, thereby preserving correlations between channels. The joint covariance then undergoes a Cholesky decomposition. This matrix

is used to generate a set of randomized polynomial coefficients that share the same covariances as the data-derived polynomial modes. In the end, this is an overly complicated way of simply adding a randomized mean and slope to the scanset that is polynomial-filtered away before pairmap generation anyway. As a result, the noise realizations are insensitive to the re-injection of the low-order polynomial modes.

The high-frequency modes are generated in much the same way, but are decomposed into Fourier modes, rather than polynomial modes. The data are again concatenated across half-scans within a scanset, and then filtered to remove the polynomial modes captured by the first part of the noise simulation. Next, the data are Fourier transformed, and then binned into logarithmically-spaced frequency bins. In this way, the noise simulator preserves correlations between channels, but treats each Fourier mode as independent. The final steps proceed exactly as in the case of the low-frequency modes: The Cholesky decomposition is calculated and used to generate random Fourier coefficients that have the same joint covariance as the real data. By performing an inverse discrete Fourier transform, the randomized coefficients are converted back into the time domain, creating TODs that are carried forward for later analysis.

An example of a randomly selected BICEP1 phase is plotted in Figure 4.3. The simulated noise power spectra is plotted against the real data. The noise model relies on the assumption that in any given half-scan, the detector time series is noise-dominated. This is a safe assumption: atmospheric fluctuations are typically of order $0.1 - 1$ K, whereas temperature anisotropies in the CMB are $4 - 5$ orders of magnitude lower in amplitude.

4.3.3 Spurious polarization generation

In addition to generating noise and signal simulations, the pipeline can also generate simulated spurious polarization from specific classes of instrument systematics. This aspect of the pipeline was expanded for the BICEP2 analysis to include a wider variety of potential sources of false polarization.

In the simplest case, the pipeline code can be used to generate spurious polarization that results from mismatched detector gains. Relative gain mismatch induced leakage is injected into the TODs by simply artificially adjusting the relative gain corrections for A and B such that the temperature signal does not exactly vanish in the pair difference. The simulator can either set a randomized relative gain correction for each scanset (with some specified distribution), or inject a uniform mismatch in the simulated data over the entire timeframe of the simulation.

Similarly, differential pointing can be simulated by displacing the assumed pointing centers for A and B from the common centroid. When differenced, the unpolarized signal measured from two nearby points for A and B will not difference away perfectly, resulting in temperature-to-polarization leakage. This can be simulated as a fixed displacement between A and B, or a randomized displacement that varies with time.

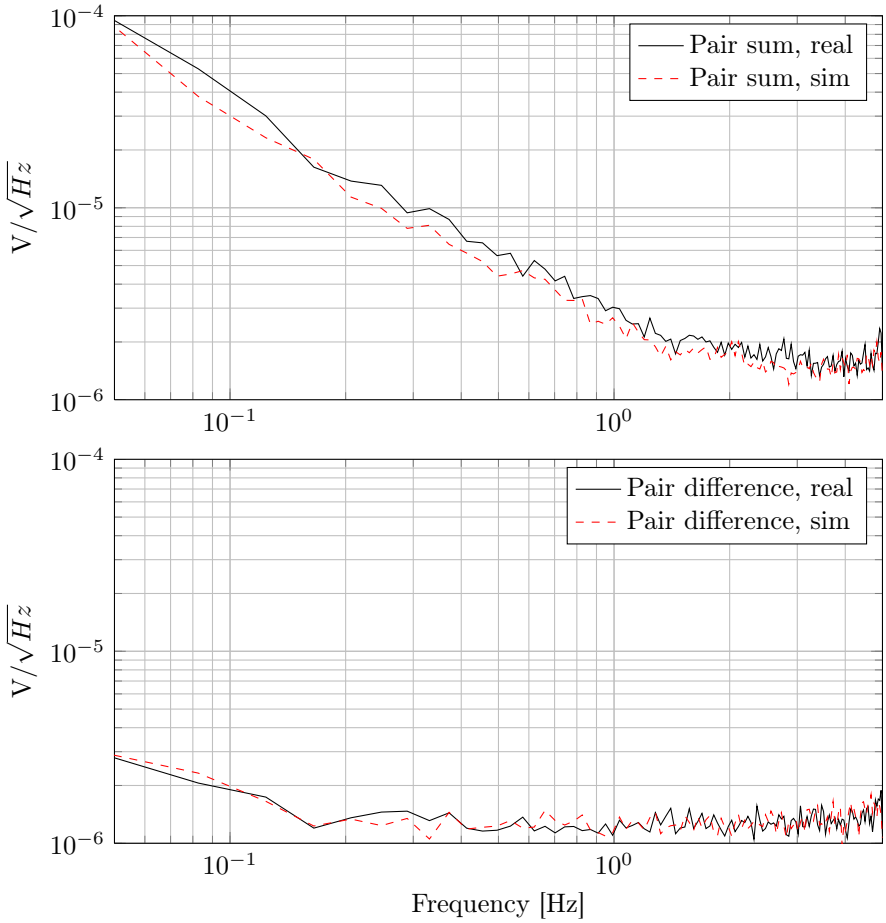


Figure 4.3: Example PSD of BICEP1 real data (black) and simulated noise (dashed red) for a randomly selected detector pair and phase. For clarity, the plotted PSD is the average across all per-half-scan spectra within a single phase. This averaging reduces sample variance, thereby making a visual comparison of the power spectra more clear.

Spurious polarization from instrument mis-calibration, such as polarization orientation uncertainty, can also be simulated by varying the assumed instrument properties. When carried through to angular power spectra, potential levels of spurious polarization can be assessed.

In the simulation pipeline reported in Takahashi et al. 2010, differential beam width and differential ellipticity were simulated by performing various beam convolutions on a flat-sky projection of the simulated curved-sky `Healpix` map. Signal-only TODs were then calculated by interpolating these flat-sky maps at locations corresponding to the detector pointing centers, as in Section 4.3.1. In the three-year analysis, as well as in the BICEP2 analysis, the interpolation algorithm was changed to sample the `Healpix` map directly, thereby skipping the intermediate flat-sky projected map. This modification required a change to the simulation procedure used to calculate differential beam width and differential ellipticity.

The BICEP2 simulation pipeline has been expanded to approximate leakage from a number of

potential sources of instrumental polarization on the curved sky, including differential ellipticity and differential beam width. These are in active development, so a finalized description of the method is not yet possible. We will, however, give a general description of the procedure. In all cases, spurious polarization is calculated by taking a linear combination of external template maps, taken as curved-sky `Healpix` temperature maps, and injecting some fraction of that linear combination into the TODs. Simulations of differential beam width, for instance, use a combination of `Healpix` maps smoothed to different effective beam widths. The signal in A and B is calculated by taking different linear combinations of the pre-smoothed maps to simulate the signal appropriate for the beam widths for σ_A and σ_B, respectively. Ellipticity is simulated in much the same way, using maps that are both smoothed to some nominal beam width and also symmetrically displaced about the common centroid. Multiple circular Gaussians with displaced centroids are used to approximate an elliptical beam. This additional capability of the BICEP2 pipeline will be required to accurately assess potential sources of systematics resulting from beam mismatch.

4.4 Instrumental polarization deprojection

In the BICEP1 two-year analysis, it was noted that two potential sources of instrumental polarization could contribute systematics relevant to bandpowers equivalent to $r = 0.1$. These two sources of systematics were relative gain mismatch (mismatch in relative gains between detectors within a polarized pair) and differential pointing (a pointing offset between detectors within a polarized pair). We have implemented new analysis techniques to mitigate these potential sources of systematics, and, as a result, we are able to place much tighter constraints on the level of potential systematic contamination of BB bandpowers.

There are two general approaches to contend with these potential systematics that may be considered. If various instrument parameters corresponding to specific sources of instrumental polarization are stable over time and measured with high signal-to-noise, then the leakage can be precisely calculated and subtracted using some high signal-to-noise measurement of the temperature anisotropy. In the event that the instrument parameters are time variable or not precisely measured, this can lead to a leakage over- or under-subtraction, resulting in a systematic error that cannot be measured with simulations. The second approach is to generate a template of the spurious polarization, but to allow the amplitude of the leakage to float as a free parameter. This results in some additional loss of information (because more degrees of freedom are allowed), but bandpowers will only ever be suppressed by an amount that is precisely calculable via simulations.

This work follows on similar analyses to estimate instrument-induced systematics, such as Hu et al. 2003, O'Dea et al. 2007, (Shimon et al. 2008), and Yadav et al. 2010. These efforts largely focused on finding analytic methods to generate predictions of spurious polarization from measured

instrument parameters. The deprojection analysis presented here is unique among these efforts in using similar estimation techniques to suppress potential sources of systematics in observational data.

For each pair, we construct a pair-difference time series $d_f(t)$ by subtracting the el-nod calibrated time series for A and B (as in Equation A.7). We seek a minimum variance estimate of α, the coefficient in the linear equation:

$$d_f(t) = \alpha \mathbf{T} + d_p(t) + n_d(t). \tag{4.16}$$

Here, $d_p(t)$ is the pair-difference signal that results from the true sky polarization (as in Equation A.13), $n_d(t)$ is the pair-difference noise, and $\alpha\mathbf{T}$ is the term that represents spurious polarization. The template $\mathbf{T}$ has dimensions of $n_{\rm samp} \times m$, where $n_{\rm samp}$ is the number of time samples and m is the number of potential sources of leakage, and α has dimensions of m. The template can thus incorporate any number of potential sources of instrument-induced polarization, including relative gain mismatch, beam mismatch, etc. We float the amplitude of these sources of leakage by performing a χ^2 minimization for α. We call α the "regression coefficients" derived from the data.

In this section we will describe the details of this deprojection algorithm, whereby we generate a template for various sources of polarization contamination (construct $\mathbf{T}$), regress the template against the data (estimate α, denoted as $\hat{\alpha}$), and then subtract it away (remove $\hat{\alpha}\mathbf{T}$). The classes of sources of spurious polarization include relative responsivity mismatch, differential pointing, differential beam width, and differential ellipticity. We begin with the pair-difference time series construction, as presented in Appendix A.1, but now allowing the relative responsivities, centroids, beam widths, and ellipticities to vary between detectors A and B (we assume idealized polarization efficiency and orthogonality):

$$\begin{aligned} d_f(t) =& \frac{1}{2}\int d\vec{p}\,'[\tilde{g}_A B_A(\vec{p}_A(t) - \vec{p}\,') - \tilde{g}_B B_B(\vec{p}_A(t) - \vec{p}\,')]\Theta(\vec{p}\,') \\ &+ \frac{1}{2}\int d\vec{p}\,'[\tilde{g}_A B_A(\vec{p}_A(t) - \vec{p}\,') + \tilde{g}_B B_B(\vec{p}_A(t) - \vec{p}\,')](Q(\vec{p}\,')\cos 2\psi(t) + U(\vec{p}\,')\sin 2\psi(t))] + n_f(t). \end{aligned} \tag{4.17}$$

Here, $\tilde{g}$ is the el-nod corrected responsivity (in ADU/K), $\vec{p}_i(t)$ is the detector pointing on the sphere[6] for detector i, $B_i(\vec{p})$ is the "as observed" frequency-independent beam function as defined in Equation A.3, $n_f(t)$ is the pair-difference noise, $\psi(t)$ is the polarization orientation angle of the pair (equal to detector A's orientation angle and assuming B is perfectly orthogonal), and Θ, Q, and U are the temperature and polarization anisotropy fields.

[6]This may be in either celestial or equatorial coordinates.

4.4.1 Relative gain mismatch deprojection

We begin with relative gain mismatch. Let us consider two detectors within a polarized pair with perfectly matched beam shapes (letting $B_A(\vec{p}) = B_B(\vec{p}) = B(\vec{p})$) and beam centroids (letting $\vec{p}_A(t) = \vec{p}_B(t) = \vec{p}_0(t)$), but mismatched in relative gain. We parameterize the relative gain mismatch as:

$$\delta\tilde{g} = \frac{\tilde{g}_A - \tilde{g}_B}{2}. \tag{4.18}$$

As before, we assume perfect polarization orthogonality and efficiency, setting $\psi_A = \psi$, $\psi_B = \psi + \pi/2$, and $\gamma = 1$ (where γ is the polarization efficiency as defined in Equation 4.4). The pair-difference time series is then constructed (re-writing the integral in Equation 4.17 as a convolution) as:

$$d_f(t) = \delta\tilde{g}(B * \Theta)(\vec{p}_0(t)) + d_p(t) + n_d(t). \tag{4.19}$$

Here, $d_p(t)$ is the pair-difference signal from the true sky polarization:

$$d_p(t) = \tilde{g}[B * (Q\cos 2\psi + U \sin 2\psi)](\vec{p}_0(t)). \tag{4.20}$$

We thus find that the leakage in the pair-difference time series is simply proportional to the beam-convolved temperature map, $(B * \Theta)(\vec{p}_0(t))$.

Because we have available high signal-to-noise measurements of the observed temperature field, we can construct a template of the leakage and regress it against the pair-difference data. The BICEP1 three-year analysis made use of the WMAP seven-year V-band (70 GHz) maps, but future analyses will be able to take advantage of higher signal-to-noise measurements (from Planck, for instance). The template for the relative gain mismatch leakage is calculated in the same manner as signal-only TODs, as described in Section 4.3.1, substituting a random realization of the CMB with a `Healpix` map from observations (WMAP V-band, in the case of BICEP1). The relative responsivitiy mismatch leakage time series template $\mathrm{T}_{\mathrm{rg}}(t)$ can be calculated as:

$$\mathrm{T}_{\mathrm{rg}}(t) = \tilde{g}_{\mathrm{abs}}(B * \Theta)(\vec{p}_0(t)). \tag{4.21}$$

Here, $\tilde{g}_{\mathrm{abs}}$ is the absolute calibration factor, as calculated in Section 4.1.5. In practice, there are time series-level filtering and weighting operations not captured in this expression. However, the time series of the template signal is generated alongside the data, ensuring that the same filtering and weighting operations are applied to the template as to the real data.

After constructing the template time series, we regress it against the data. In the case of BICEP1, this is performed prior to binning the TODs into maps. The weighted, filtered template is regressed against the weighted, filtered data across roughly 10 hours of data simultaneously. We calculate a

minimum variance estimate of α using a weighted linear regression:

$$\hat{\alpha}_{\rm rg} = [{\rm T_{rg}}^{\top}\mathcal{N}^{-1}{\rm T_{rg}}]^{-1}{\rm T_{rg}}^{\top}\mathcal{N}^{-1}d_f. \tag{4.22}$$

This is analogous to the standard mapmaking equation. In the case of BICEP1, the weight matrix $\mathcal{N}^{-1}$ is constructed from the inverse variance of each half-scan within a scanset, thereby down-weighting periods of noisy or aberrant data.

The true value of the relative-responsivity mismatch regression coefficient is trivially related to the relative gain mismatch:

$$\alpha_{rg} = \frac{\delta\tilde{g}}{\tilde{g}_{\rm abs}}. \tag{4.23}$$

The factor of $\tilde{g}_{\rm abs}$ in the denominator comes about because we have chosen to convert the template into the same units as the detector time series prior to regression. We also note that this expression relies on the assumption that there is no correlation between the noise or the sky polarization and the template. In the case of BICEP1, we have demonstrated that this is true to a very good approximation. While there is correlation between the temperature field (from which the template is constructed) and the E-modes, the correlation is both positive and negative across the angular scales at which we observe. As a result, for any given sweep of a particular detector across our observing region, the map-space correlation will have a randomized amplitude and sign. Averaged over many detectors and DK angles, we empirically find that the correlation averages to near zero. This may not be the case for small angular scale experiments targeting angular scales for which the TE correlation is always positive. However, this correlation can be accounted for in simulations.

The choice to perform the regression at the TOD level is largely one of convenience: In the BICEP2 analysis, the templates are first binned into pairwise maps (as described in Section 4.1.3) before regressing. With the data-derived regression coefficients in-hand, we can readily subtract away the template multiplied by the regression coefficient to "clean" the data of relative gain mismatch leakage. We will explore this step in detail in Chapter 5. We can also use the regression coefficients as inputs to simulations to predict the level of spurious polarization prior to implementing the cleaning step of the deprojection analysis.

We can use this approach to combat spurious polarization from beam mismatch in much the same way. In the following sections, the same generalized approach is taken, but different template constructions are required.

4.4.2 Differential pointing deprojection

Next, we consider temperature-to-polarization leakage resulting from two detectors on the sky with idealized beam shapes, gain corrections, and polarization properties, but displaced on the sky by some small angular distance. We let $B(\vec{p}_i(t))$ represent the beam as a function of the time-dependent detector pointing, $\vec{p}_i(t)$, which we allow to vary between A and B. We define the average beam pointing of the two detectors, as well as a symmetric displacement of A and B, from the common centroid:

$$\vec{p}_o(t) = \frac{\vec{p}_A(t) + \vec{p}_B(t)}{2} \tag{4.24}$$

$$\overrightarrow{\delta p} = \frac{\vec{p}_A(t) - \vec{p}_B(t)}{2}. \tag{4.25}$$

The beam function can be approximated at some nearby location to the common centroid via a simple Taylor series expansion. For detector A, this is simply:

$$B(\vec{p}_0(t) + \overrightarrow{\delta p}) = B(\vec{p}_0(t)) + \overrightarrow{\delta p} \cdot \nabla B(\vec{p})|_{\vec{p}=\vec{p}_0(t)} + \mathcal{O}(|\overrightarrow{\delta p}|)^2. \tag{4.26}$$

When we form the pair-difference beam, the zeroth-order term cancels, along with even powers of the expansion variable, leaving:

$$B(\vec{p}_0(t) + \overrightarrow{\delta p}) - B(\vec{p}_0(t) - \overrightarrow{\delta p}) = 2\overrightarrow{\delta p} \cdot \nabla B(\vec{p})|_{\vec{p}=\vec{p}_0(t)} + \mathcal{O}(|\overrightarrow{\delta p}|)^3. \tag{4.27}$$

In celestial coordinates, the displacement vector $\overrightarrow{\delta p}$ changes with the boresight rotation of the telescope. It is convenient to parameterize the displacement in coordinates that are invariant under boresight rotation. If we let $\hat{\phi}$ and $\hat{\theta}$ be "physics" spherical unit vectors,[7] then we can define "focal plane coordinates" $\hat{x}$ and $\hat{y}$ as a function of the boresight rotation angle DK:

$$\hat{x} = -\hat{\phi}\cos DK + \hat{\theta}\sin DK,\ \hat{y} = -\hat{\phi}\sin DK - \hat{\theta}\cos DK. \tag{4.28}$$

Note that a similar but different Cartesian coordinate system is used in Section 3.1.2 to define various beam parameters, and the two are not to be confused. We use this convention simply as a convenient intermediate step for calculating the leakage. Note that $\hat{\phi}$ and $\hat{\theta}$ have unit length. The

[7] The usual physics θ and ϕ are related to declination and right ascension as $\theta = 90-\text{DEC}$ and $\phi = \text{RA}$.

gradients in the $\hat{x}$ and $\hat{y}$ direction in spherical coordinates are:

$$\begin{aligned}\hat{x}\cdot\nabla &= \left[-\hat{\phi}\cos DK + \hat{\theta}\sin DK\right]\cdot\left[\hat{\phi}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} + \hat{\theta}\frac{\partial}{\partial\theta}\right] \\ &= -\frac{\cos DK}{\sin\theta}\frac{\partial}{\partial\phi} + \sin DK\frac{\partial}{\partial\theta} \qquad (4.29)\\ \hat{y}\cdot\nabla &= \left[-\hat{\phi}\sin DK - \hat{\theta}\cos DK\right]\cdot\left[\hat{\phi}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} + \hat{\theta}\frac{\partial}{\partial\theta}\right] \\ &= -\frac{\sin DK}{\sin\theta}\frac{\partial}{\partial\phi} - \cos DK\frac{\partial}{\partial\theta}. \qquad (4.30)\end{aligned}$$

We now consider the specific case where A and B are displaced purely along the $\hat{x}$ direction by a magnitude equal to $\delta x = \vec{\delta p}\cdot\hat{x}$.[8] The pair-difference beam is thus:

$$B(\vec{p}_0(t)+\delta x\hat{x}) - B(\vec{p}_0(t)-\delta x\hat{x}) = 2\delta x\left(-\frac{\cos DK}{\sin\theta}\frac{\partial}{\partial\phi} + \sin DK\frac{\partial}{\partial\theta}\right)B(\vec{p})|_{\vec{p}=\vec{p}_0(t)} + \mathcal{O}(\delta x)^3. \qquad (4.31)$$

Returning to the temperature-to-polarization leakage, we can write the resulting pair-sum and pair-difference time series by inserting our expression for the pair-difference beam into Equation 4.17 (assuming idealized responsivity corrections):

$$\begin{aligned}d_f(t) \approx\ & \tilde{g}\delta x\int d\vec{p}'\left(-\frac{\cos DK}{\sin\theta}\frac{\partial}{\partial\phi} + \sin DK\frac{\partial}{\partial\theta}\right)B(\vec{p})|_{\vec{p}=\vec{p}_0(t)-\vec{p}'}\Theta(\vec{p}') \\ &+ \tilde{g}\int d\vec{p}'B(\vec{p}_0(t)-\vec{p}')[Q(\vec{p}')\cos 2\psi(t) + U(\vec{p}')\sin 2\psi(t)] + n_f(t). \qquad (4.32)\end{aligned}$$

Note that in the last step, we have ignored distortions to the polarization field due to the mismatched beam centroids. This distortion, proportional to the gradient of Q and U, will be smaller than the temperature-to-polarization leakage by several orders of magnitude, due to the CMB's small fractional polarization. We can rewrite the differential signal as a convolution, and for clarity let $d_p(t)$ represent the undistorted pair-difference signal from the true sky polarization:

$$d_f(t) \approx \tilde{g}\delta x\left(-\frac{\cos DK}{\sin\theta}\frac{\partial}{\partial\phi} + \sin DK\frac{\partial}{\partial\theta}\right)(B*\Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)} + d_p(t) + n_f(t). \qquad (4.33)$$

In this last step, we have taken advantage of a fundamental property of convolution: Taking the convolution of the derivative of the beam with the temperature field is equivalent to taking the derivative of the beam-convolved temperature field. In other words, $f'(x)*g(x) = f(x)*g'(x) = (f*g)'(x)$.

We thus arrive at what we might have initially guessed: to first order, differential pointing leakage is equal to the spatial derivative of the beam-smoothed temperature field. This is readily available:

[8] Note that this is only true for infinitesimal displacements in x.

the first and second map derivatives on the sphere are standard data products produced by the `synfast` package. In simulation, we can construct a template for a focal plane-fixed displacement δx by taking a linear combination of the map spatial derivatives on the sphere. We will denote this template as T_x (illustrated in Figure 4.4):

$$\mathrm{T}_x(t) = \tilde{g}\,\sigma_0 \left(-\frac{\cos DK}{\sin\theta}\frac{\partial}{\partial\phi} + \sin DK \frac{\partial}{\partial\theta} \right) (B * \Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)}. \tag{4.34}$$

The factor of the nominal beam width, σ_0, is added to keep all templates in common units (ADU). We can estimate the corresponding regression coefficient $\hat{\alpha}_x$ as in Equation 4.22, but substituting in our new template. We find that the true value of α_x is, to leading order, simply related to the displacement magnitude:

$$\alpha_{\delta x} = \frac{\delta x}{\sigma_0} + \mathcal{O}(\delta x)^3. \tag{4.35}$$

We find the same to be true for an infinitesimal displacement in $\hat{y}$:

$$\mathrm{T}_y(t) = \tilde{g}\,\sigma_0 \left(-\frac{\sin DK}{\sin\theta}\frac{\partial}{\partial\phi} - \cos DK \frac{\partial}{\partial\theta} \right) (B * \Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)} \tag{4.36}$$

$$\alpha_{\delta y} = \frac{\delta y}{\sigma_0} + \mathcal{O}(\delta y)^3. \tag{4.37}$$

4.4.3 Differential beam width deprojection

In order to calculate the leakage from differential beam width, we now need to invoke a specific beam model of a two-dimensional elliptical Gaussian. For concreteness, we assume a beam center $\bar{p}$, a vector on the sphere that points from the origin to the maximum in the beam response. We assume that $\vec{p}$ and $\bar{p}$ are defined in the same coordinate system on the sphere. It is convenient to describe the beam in a flat-sky projection. To do so, we define a new coordinate $\vec{x}$ as:

$$\vec{x} = [(\vec{p} - \bar{p}) \cdot \hat{x}]\hat{x} + [(\vec{p} - \bar{p}) \cdot \hat{y}]\hat{y}. \tag{4.38}$$

Here, $\hat{x}$ and $\hat{y}$ are as defined in Equation 4.28 and $\bar{p}$ is the beam centroid. Very near the beam centroid ($\vec{x} = [0, 0]$), the beam is well-described by a flat-sky projection. In the flat-sky limit, the beam function can be approximated as:

$$B(\vec{x}) = \frac{1}{\Omega} \exp\left[-\frac{1}{2}\vec{x}^{\top}\Sigma^{-1}\vec{x} \right]. \tag{4.39}$$

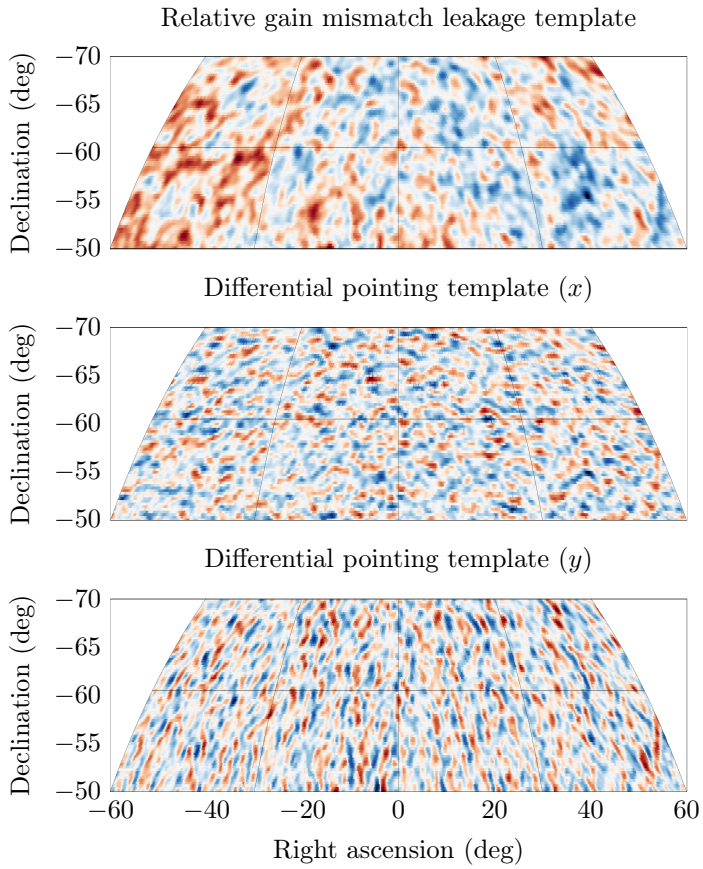


Figure 4.4: *Top Panel* : The map used for constructing the relative gain mismatch contamination leakage. This comes from simply re-sampling the WMAP V-band map convolved with a Gaussian-approximated BICEP1 beam. *Middle Panel* : The map template for differential pointing leakage from an infinitesimal displacement in x. This is a linear combination of the first spatial derivatives of the top map in θ and ϕ, as calculated by `synfast`. *Bottom Panel* : The map template for differential pointing leakage from an infinitesimal displacement in y.

Here, Σ is a 2×2 matrix that describes the beam shape, in analogy with the covariance matrix for a two-dimensional Gaussian.

Now we can treat the specific case of differential beam width: We consider co-located beams for A and B that differ in their beam width, σ_A and σ_B. As before, we assume ideally corrected relative responsivities and idealized polarization properties. This gives the beam function:

$$B(\vec{x};\sigma) = \frac{1}{\Omega} \exp\left[-\frac{1}{2}\vec{x}^{\top}\Sigma_{\sigma}^{-1}\vec{x}\right]. \tag{4.40}$$

For clarity, we have written B as an explicit function of σ. Here Σ_σ is the beam covariance of a circular Gaussian, $\Sigma_\sigma = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. For a circularly symmetric Gaussian beam, the beam solid angle is calculated as $\Omega = 2\pi\sigma^2$. We define $\delta\sigma$ and σ_0, the pair-difference and

pair-average beam widths:

$$\sigma_0 = \frac{\sigma_A + \sigma_B}{2} \tag{4.41}$$

$$\delta\sigma = \frac{\sigma_A - \sigma_B}{2}. \tag{4.42}$$

We can write the beam function for detector A as an explicit function of $x = |(\vec{p} - \bar{p}) \cdot \hat{x}|$ and $y = |(\vec{p} - \bar{p}) \cdot \hat{y}|$:

$$B(\vec{x}; \sigma_A = \sigma_0 + \delta\sigma) = \frac{1}{2\pi(\sigma_0 + \delta\sigma)^2} \exp\left[-\frac{x^2 + y^2}{2(\sigma_0 + \delta\sigma)^2}\right]. \tag{4.43}$$

We can approximate $B(\vec{x}; \sigma_0 + \delta\sigma)$ as an expansion of the beam function about the pair-average beam width σ_0 in powers of $\delta\sigma$ as:

$$B(\vec{x}; \sigma_0 + \delta\sigma) \approx B_0(\vec{x}; \sigma_0) + \delta\sigma \frac{\partial}{\partial\sigma} B(\vec{x}; \sigma)|_{\sigma=\sigma_0} + \mathcal{O}(\delta\sigma)^2. \tag{4.44}$$

We now consider the differenced beam formed by calculating $B(\vec{x}; \sigma_0 + \delta\sigma) - B(\vec{x}; \sigma_0 - \delta\sigma)$. The zeroth-order terms cancel, as well as even powers of $\delta\sigma$:

$$B(\vec{x}; \sigma_0 + \delta\sigma) - B(\vec{x}; \sigma_0 - \delta\sigma) = 2\delta\sigma \frac{\partial}{\partial\sigma} B(\vec{x}; \sigma)|_{\sigma=\sigma_0} + \mathcal{O}(\delta\sigma)^3. \tag{4.45}$$

We can write the partial derivative with respect to σ explicitly in terms of x and y:

$$\frac{\partial}{\partial\sigma} B(\vec{x}; \sigma) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^3} B(\vec{x}; \sigma). \tag{4.46}$$

This can be rewritten in terms of second-order partial derivatives:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) B(\vec{x}; \sigma) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} B(\vec{x}; \sigma) = \frac{1}{\sigma}\frac{\partial}{\partial\sigma} B(\vec{x}; \sigma). \tag{4.47}$$

We now move back to spherical coordinates. We invoke a flat-sky approximation by assuming that the description of the beam in x and y is equivalent to the beam function on the sphere. Taking the definitions of $\hat{x}$ and $\hat{y}$, we can re-write the second partials in terms of spherical coordinates $\vec{p} = \{\theta, \phi\}$ as:

$$\frac{\partial^2}{\partial x^2} = \cos DK \sin DK \left(\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \sin^2 DK \frac{\partial^2}{\partial\theta^2} + \frac{\cos^2 DK}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{4.48}$$

$$\frac{\partial^2}{\partial y^2} = -\cos DK \sin DK \left(\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \cos^2 DK \frac{\partial^2}{\partial\theta^2} + \frac{\sin^2 DK}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{4.49}$$

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \tag{4.50}$$

We can thus re-express the difference beam (up to third order in $\delta\sigma$) as:

$$B(\vec{p};\sigma_0+\delta\sigma)-B(\vec{p};\sigma_0-\delta\sigma)\approx 2\delta\sigma\sigma_0\left(\frac{\partial^2}{\partial\theta^2}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)B(\vec{p};\sigma)|_{\sigma=\sigma_0}. \tag{4.51}$$

The fact that there is no DK dependence makes intuitive sense, given the rotational symmetry of the problem. For this reason, differential beam width leakage is sometimes called "monopole" leakage.

Using this rather simple expression, we can return to the temperature-to-polarization leakage in the pair-difference time series. As in the case of differential pointing, we can express it in terms of derivatives on the sphere. The total pair-difference signal seen by the detector pair is:

$$d_f(t)\approx 2(\delta\sigma\sigma_0)\tilde{g}\left(\frac{\partial^2}{\partial\theta^2}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)(B*\Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)}+d_p(t)+n_f(t). \tag{4.52}$$

Here, $(B*\Theta)(\vec{p}(t))$ is the observed time-ordered beam-convolved temperature map. As in the case of differential pointing, we have expressed the leakage in terms of the derivatives of the convolved map. The explicit dependence on the beam width has been suppressed, since the beam width is now fixed at σ_0. In simulation, the first- and second-order derivatives in θ and ϕ of the beam-convolved temperature map are readily available from `synfast`. We use these to construct the following template for differential beam width:

$$\mathrm{T}_{\delta\sigma}=2\tilde{g}\sigma_0^2\left(\frac{\partial^2}{\partial\theta^2}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)(B*\Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)}. \tag{4.53}$$

As in previous sections, we regress the template against the data to arrive at a best estimate of $\alpha_{\delta\sigma}$. The true value of $\alpha_{\delta\sigma}$ is simply related to the differential beam width as:

$$\alpha_{\delta\sigma}\approx\frac{\delta\sigma}{\sigma_0}+\mathcal{O}(\delta\sigma/\sigma)^3. \tag{4.54}$$

4.4.4 Differential ellipticity deprojection

We now consider mismatch in the beam ellipticities of A and B. Unlike the previous cases, symmetrizing the displacement about the mean of A and B is not particularly meaningful, as we care about the overall magnitude of the beam ellipticity as well as the mismatch between detectors. For our purposes, however, we will assume that the beam ellipticity of the pair-sum is small. This is a good approximation for most pixels on the focal plane. Aberrant pixels with large ellipticities (of which there are only a few) can always be cut from analysis.

We first consider ellipticity in the 'plus' orientation (co-aligned with $\hat{x}$ and $\hat{y}$). The beam covari-

ance is described as:

$$\Sigma_p = \begin{bmatrix} \sigma_0^2(1+p) & 0 \\ 0 & \sigma_0^2(1-p) \end{bmatrix}. \tag{4.55}$$

We assume ideally matched corrected gains, beam centroids, and beam widths, setting the beam width to σ_0. The beam function, written as an explicit function of the ellipticity parameter, is thus:

$$B(\vec{x};p) = \frac{1}{2\pi\sigma_0^2(1-p^2)} \exp\left[-\frac{1}{2}\left(\frac{x^2}{2\sigma_0^2(1+p)} + \frac{y^2}{2\sigma_0^2(1-p)}\right)\right]. \tag{4.56}$$

Expanding in powers of p, we find:

$$B(\vec{x};p) \approx B(\vec{x};p=0) + p\frac{\partial}{\partial p}B(\vec{x};p)|_{p=0} + \mathcal{O}(p)^2. \tag{4.57}$$

The partial derivative with respect to p can be rewritten in terms of the second partials in x and y:

$$\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}\right) B(\vec{x};p=0) = \frac{x^2-y^2}{\sigma_0^4} B(\vec{x};p=0) = \frac{1}{\sigma_0^2}\frac{\partial}{\partial p}B(\vec{x};p)|_{p=0}. \tag{4.58}$$

As before, we can translate derivatives in x and y into spherical coordinates:

$$\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} = \sin 2DK\left(\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \cos 2DK\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right). \tag{4.59}$$

With these spherical derivatives in hand, we can calculate the differenced beam in the presence of differential ellipticity:

$$\begin{aligned} &B(\vec{p};p_A) - B(\vec{p};p_B) = \\ &(p_A - p_B)\sigma_0^2\left[\sin 2DK\left(\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \cos 2DK\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right)\right] B(\vec{p};p=0) \\ &+ \mathcal{O}(p_A^2 - p_B^2). \end{aligned} \tag{4.60}$$

Note the 180-degree rotational symmetry in DK. Differential ellipticity has spin-2 symmetry on the sky, and for this reason is sometimes called "quadrupole" leakage. This is consistent with our illustration in Figure 3.5.

We use Equation 4.60 to calculate the leakage from a small differential ellipticity in A and B:

$$d_f(t) = \frac{\tilde{g}}{2}(B(\vec{p}_0(t);p_A) - B(\vec{p}_0(t);p_B)) * \Theta(\vec{p}_0(t)) + d_p(t) + n_f(t) \tag{4.61}$$

$$\begin{aligned} &\approx \tilde{g}\,\delta p\,\sigma_0^2\left[\sin 2DK\left(\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \cos 2DK\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right)\right](B*\Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)} \\ &+ d_p(t) + n_f(t). \end{aligned} \tag{4.62}$$

Here, $\delta p = (p_A - p_B)/2$ is the unitless differential beam ellipticity in the 'plus' orientation.

We simulate differential ellipticity by taking a DK-dependent linear combination of the partial derivatives on the sphere as computed by `synfast`. The template is constructed as:

$$\mathrm{T}_p = \tilde{g}\,\sigma_0^2 \left[\sin 2DK \left(\frac{\cos\theta}{\sin^2\theta}\frac{\partial}{\partial\phi} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \cos 2DK \left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right)\right] (B * \Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)}. \tag{4.63}$$

The corresponding regression coefficient, α_p is then:

$$\alpha_p \approx \delta p + \mathcal{O}(p_A^2 - p_B^2). \tag{4.64}$$

Note that unlike the cases of differential pointing and differential ellipticity, there is no way to symmetrize the ellipticity about the pair-average. In the event that the individual ellipticities are large, there will be non-negligible temperature-to-polarization leakage, even if the differential ellipticity is small. Our approximation ignores this effect and is thus a poor approximation when the individual detectors' beam ellipticities are large.

Differential mode	Symbol	Definition	Regression coefficient	Template
Relative gain mismatch	$\delta\tilde{g}$	$(\tilde{g}_A - \tilde{g}_B)/2$	$\alpha_{rg} = \delta\tilde{g}/\tilde{g}_{\mathrm{abs}}$	Equation 4.21
Differential pointing, x	δx	$\vec{\delta p}\cdot\hat{x}$	$\alpha_{\delta x} = \delta x/\sigma_0$	Equation 4.34
Differential pointing, y	δy	$\vec{\delta p}\cdot\hat{y}$	$\alpha_{\delta y} = \delta y/\sigma_0$	Equation 4.36
Differential beam width	$\delta\sigma$	$(\sigma_A - \sigma_B)/2$	$\alpha_{\delta\sigma} = \delta\sigma/\sigma_0$	Equation 4.53
Differential ellipticity, +	δp	$(\delta p_A - \delta p_B)/2$	$\alpha_p = \delta p$	Equation 4.71
Differential ellipticity, ×	δc	$(\delta c_A - \delta c_B)/2$	$\alpha_c = \delta c$	Equation 4.71

Table 4.3: Differential parameters and regression coefficients. The el-nod corrected responsivity $\tilde{g}$ is defined in Equation A.9. The differential pointing vector, $\vec{\delta p}$ is defined in Equation 4.25.

The final deprojection case we consider is for differential ellipticity in the 'cross' direction parameterized by c:

$$\Sigma_c = \begin{bmatrix} \sigma_0^2 & c\sigma_0^2 \\ c\sigma_0^2 & \sigma_0^2 \end{bmatrix}. \tag{4.65}$$

We once again adjust our notation to write the flat-sky approximated beam function as an explicit function of the cross ellipticity, $B(\vec{x}; c)$. Expanding in terms of c, we find:

$$B(\vec{x}; c) \approx B(\vec{x}; c = 0) + c\frac{\partial}{\partial c}B(\vec{x}; c)|_{c=0} + \mathcal{O}(c)^2. \tag{4.66}$$

As before, we can express the partial derivative with respect to c in terms of spatial derivatives:

$$\left(\frac{\partial^2}{\partial y \partial x} + \frac{\partial^2}{\partial x \partial y}\right) B(\vec{x}; c=0) = \frac{2xy}{\sigma_0^4} B(\vec{x}; c=0) = -\frac{2}{\sigma_0^2}\frac{\partial}{\partial c} B(\vec{x}; c)|_{c=0}. \tag{4.67}$$

In spherical coordinates, this becomes:

$$\frac{\partial^2}{\partial y \partial x} + \frac{\partial^2}{\partial x \partial y} = -\cos 2DK \left(\frac{\cos\theta}{\sin^2\theta} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \sin 2DK \left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right). \tag{4.68}$$

Note that this is exactly what we would get from Equation 4.59 were we to rotate DK by -45 degrees. Perturbations to c thus look exactly like perturbations to p, but rotated by 45 degrees. We can construct a pair-difference beam as:

$$\begin{aligned} &B(\vec{p}; c_A) - B(\vec{p}; c_B) = \\ &\frac{c_A - c_B}{2}\sigma_0^2 \left[-\cos 2DK \left(\frac{\cos\theta}{\sin^2\theta} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \sin 2DK \left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right)\right] B(\vec{p}; c=0) \\ &+ \mathcal{O}(c_A^2 - c_B^2). \end{aligned} \tag{4.69}$$

Plugging this expression into our calculation of the pair-difference time series and defining $\delta c = (c_A - c_B)/2$, we find:

$$\begin{aligned} d_f(t) =& \frac{\tilde{g}}{2}\delta c\, \sigma_0^2 \left[-\cos 2DK \left(\frac{\cos\theta}{\sin^2\theta} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \sin 2DK \left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right)\right] (B * \Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)} \\ &+ d_p(t) + n_f(t). \end{aligned} \tag{4.70}$$

The template for differential ellipticity in the cross direction is constructed from the second derivatives of the `synfast` maps as:

$$\mathrm{T}_c = \frac{\tilde{g}}{2}\sigma_0^2 \left[-\cos 2DK \left(\frac{\cos\theta}{\sin^2\theta} - \frac{1}{\sin\theta}\frac{\partial^2}{\partial\phi\partial\theta}\right) + \sin 2DK \left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial\theta^2}\right)\right] (B * \Theta)(\vec{p})|_{\vec{p}=\vec{p}_0(t)}. \tag{4.71}$$

Regressing the template against $d_f(t)$, we find the true value of the regression coefficient α_c will be:

$$\alpha_c \approx \delta c + \mathcal{O}(c_A^2 - c_B^2). \tag{4.72}$$

We have thus calculated templates and regression coefficients for six potential sources of leakage. We summarize the leakage type, template construction, and regression coefficient in Table 4.3.

4.5 Quadratic estimator construction for r

The tensor-to-scalar ratio is defined in such a way that observed B-mode bandpowers will increase monotonically with r. For this reason, we can estimate r using an optimally weighted sum of BB bandpowers, which is a quadratic function of the data. We call this a "quadratic estimate" of r, invoking the language used by Bond et al. 2000 and others. We describe this construction in detail in this section.

From the final bandpowers calculated using the procedure described in Section 4.2, we can construct a quadratic estimate directly using signal and noise simulations. (In the case of BICEP1, this is performed with the frequency-combined spectra, as calculated in Section 4.2.5). The observed (frequency combined) binned bandpowers used here are the final product of the power spectrum analysis; we have accounted for noise bias, filter suppression, and imperfect E/B separation. We seek a minimum variance estimate of ρ, where ρ is the coefficient in the linear equation:

$$\hat{\mathcal{C}}_b^{BB} = \rho \langle \mathcal{C}_b^{BB} \rangle + n. \tag{4.73}$$

Here $\hat{\mathcal{C}}_b^{BB}$ are the observed binned BB bandpowers, $\langle \mathcal{C}_b^{BB} \rangle$ is the expected value from a theory input BB spectrum (calculated as in Equation 4.11), and n is a noise term. We estimate ρ via a χ^2 minimization of $(\hat{\mathcal{C}}_b^{BB} - \rho \langle \mathcal{C}_b^{BB} \rangle)$, yielding:

$$\hat{\rho} = \frac{\langle \mathcal{C}_b^{BB} \rangle^\top \mathcal{N}^{-1} \hat{\mathcal{C}}_b^{BB}}{\langle \mathcal{C}_b^{BB} \rangle^\top \mathcal{N}^{-1} \langle \mathcal{C}_b^{BB} \rangle}. \tag{4.74}$$

If we regard the theory input spectrum as the predicted spectrum from tensor perturbations in the CMB, then ρ is directly proportional to r, the tensor-to-scalar ratio. Colloquially, we say that ρ informs us of the "r-ness" of our data. Here $\mathcal{N}$ is a covariance matrix calculated from signal-plus-noise simulations, containing E-mode polarization but no B-mode polarization[9]. The term $\langle \mathcal{C}_b^{BB} \rangle$, appearing in Equation 4.74, introduces a model-dependent weighting of the data. Contrastingly, since the covariance matrix is calculated from simulations containing no B-mode power, $\mathcal{N}$ is independent of the shape of the theory BB spectrum. The bandpower weight vector $\langle \mathcal{C}_b^{BB} \rangle^\top \mathcal{N}^{-1}$, calculated for both the BICEP1 frequency-combined bandpowers and individual frequencies, is reported in Table 4.4 (as reported internally in Bischoff et al. 2013). We note that the first two bins at $\ell = 37.5$ and $\ell = 72.5$ contribute over 80% of BICEP1's constraining power on r.

We calculate $\hat{\rho}$ for the real data and for each realization within our simulation set. We repeat this exercise for additional simulations that contain cosmological B-mode power. We then vary the amplitude of the B-mode spectrum to span a wide range of potential values of r, which we will call

[9] We do not include B-mode polarization from lensing in the calculation of $\mathcal{N}$. This is due to the fact that BICEP has very little sensitivity to B-mode polarization from lensing because of the beam roll-off.

Bin	ℓ range	100-auto weight	150-auto weight	Cross weight	Frequency combined
2	$20-55$	0.063	0.125	0.187	0.375
3	$55-90$	0.062	0.168	0.211	0.441
4	$90-125$	0.015	0.061	0.073	0.149
5	$125-160$	3.5×10^{-4}	0.021	0.008	0.029
6	$160-195$	-1.0×10^{-4}	0.002	0.002	0.004
7	$195-230$	8.5×10^{-5}	4.6×10^{-4}	4.6×10^{-4}	0.001
8	$230-265$	1.6×10^{-5}	3.9×10^{-4}	1.1×10^{-4}	5×10^{-4}
9	$265-300$	1.7×10^{-6}	1.5×10^{-4}	1.7×10^{-5}	2×10^{-4}
10	$300-335$	-2.3×10^{-7}	1.6×10^{-5}	4.9×10^{-6}	$<1\times10^{-4}$

Table 4.4: Relative bandpower weights for the quadratic estimate of r, taken from Bischoff et al. 2013. The frequency-combined weights are the values that are actually used to form a weighted combination of bandpowers, while the individual frequency components ($100\times100, 150\times150, 100\times150$) are provided for comparison only. Negative weights are due to random negative correlations, but are not present in the frequency-combined weights.

$r_{\rm sim}$. For ease of calculation, rather than computing independent realizations of the theory spectrum $\mathcal{C}_\ell$ for each incremented value of $r_{\rm sim}$, we generate 499 realizations of B-no-E signal-only maps for $r_{\rm sim}=0.1$. We then scale the amplitude of those same B-mode maps for $r_{\rm sim}=0$ to $r_{\rm sim}=5$, in increments of $\Delta r_{\rm sim}=0.001$. This enormous computational shortcut results in bandpowers that are independent for any given value of $r_{\rm sim}$, but are not independent across $r_{\rm sim}$.

The scaled maps are collapsed into power spectra, frequency combined (in the case of BICEP1), and then further collapsed into values of $\hat{\rho}$. This results in a 499×5000 matrix of simulated values of $\hat{\rho}$ (499 realizations at each value of $r_{\rm sim}$, 5000 values of $r_{\rm sim}$). Since the real and simulated bandpowers themselves are unbiased, the $\hat{\rho}$ estimate derived from simulations is also unbiased. There is a simple linear scaling required to translate $\hat{\rho}$ into an equivalent value of r. This is calculated by finding the scaling between $r_{\rm sim}$ and the mean value of $\hat{\rho}$ across all realizations at fixed $r_{\rm sim}$, such that $\alpha\langle\hat{\rho}\rangle = r_{\rm sim}$. This linear relation can then be applied to all realizations of $\hat{\rho}$ to compute outcomes $\hat{r}$, calculated as $\hat{r}=\alpha\hat{\rho}$. We can also apply this scaling to the real data and arrive at a quadratic estimate of r.

The ensemble of simulated outcomes of $\hat{\rho}$ can also be used to numerically evaluate parameter likelihoods. We explore this in the next section.

4.6 Probabilities, likelihoods, and posterior probability distributions

In many previous CMB analyses, an approximation to the bandpower likelihood function was calculated using the so-called "offset log-normal" approximation, first presented by Bond et al. 2000. More recently, it has become computationally feasible to evaluate the likelihood function for ρ (de-

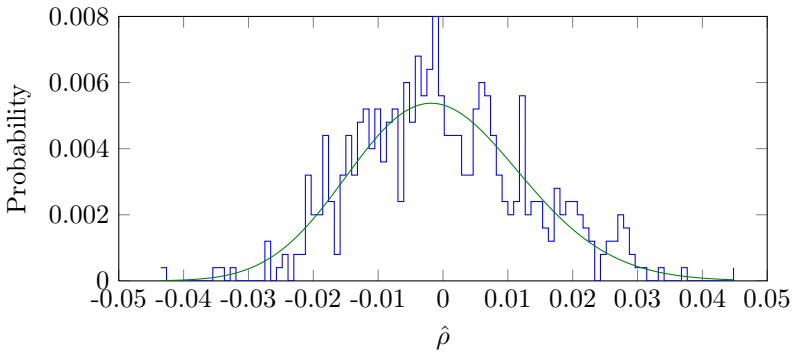


Figure 4.5: The probability $P(\hat{\rho}\,|\,r_{\rm sim}=0)$ fit to $\hat{\rho}$ outcomes for BICEP1 three-year simulations. The blue histogram represents binned outcomes of $\hat{\rho}$ for $r_{\rm sim}=0$. The green curve is the χ^2 fit calculated from the number of degrees of freedom, a scaling, and an overall offset. The probability has been normalized to unity.

fined in Equation 4.73) numerically. In this section, we review the steps whereby simulated values of ρ are used to construct likelihoods and posterior probability distributions for r.

To begin, we introduce a bit of notation. As before, $r_{\rm sim}$ is the input value to simulations. Each discrete value of $r_{\rm sim}$ yields an ensemble of bandpowers, from which we calculate $\hat{\rho}_i$ for the i-th realization (as per Equation 4.74). We reserve the '0' subscript to indicate the binned bandpowers (and subsequent value of $\hat{\rho}$) calculated for the real data. As before, the 'hats' are used to distinguish a specific parameter estimate (or bandpower) from the model parameter (or theory bandpower). We apply a simple scaling to $\hat{\rho}_i$ to arrive at a quadratic estimate of r, which we will call $\hat{r}_i$.

The probability of "outcomes" of $\hat{\rho}$, given some input value to the simulation $r_{\rm sim}$, is denoted as $P(\hat{\rho}\,|\,r_{\rm sim})$. In theory, we can numerically evaluate the probability, given an infinite number of trials and an infinite number of simulated values of $r_{\rm sim}$. In practice, we have a limited ensemble of outcomes from which a smooth probability function can be derived. There are two straightforward methods for fitting a smooth function to these outcomes: The first is to fit a function to the histogrammed bandpowers. Given our finite number of realizations, in practice the fit is sensitive to the histogram bin width and to outliers. The second method is to assume some probability distribution function and simply fit the relevant parameters. We find that the distribution is well-approximated by a χ^2 distribution (including an offset and a scaling). Since $\hat{\rho}$ is the sum of bandpowers, we expect it to be well-represented by a χ^2 distribution. (The outcomes are not exactly χ^2 distributed due to the frequency cross-spectra, which is not positive-definite, and because the bandpowers are not completely uncorrelated). Empirically, we find that the χ^2 distribution accurately reflects the distribution of $\hat{\rho}$ (Figure 4.5). The fits are normalized to integrate to unity.

It is illustrative to plot the probabilities for all values of $r_{\rm sim}$, as in Figure 4.6. The input parameter $r_{\rm sim}$ is increasing along the vertical axis. A horizontal slice at a given $r_{\rm sim}$ will yield a

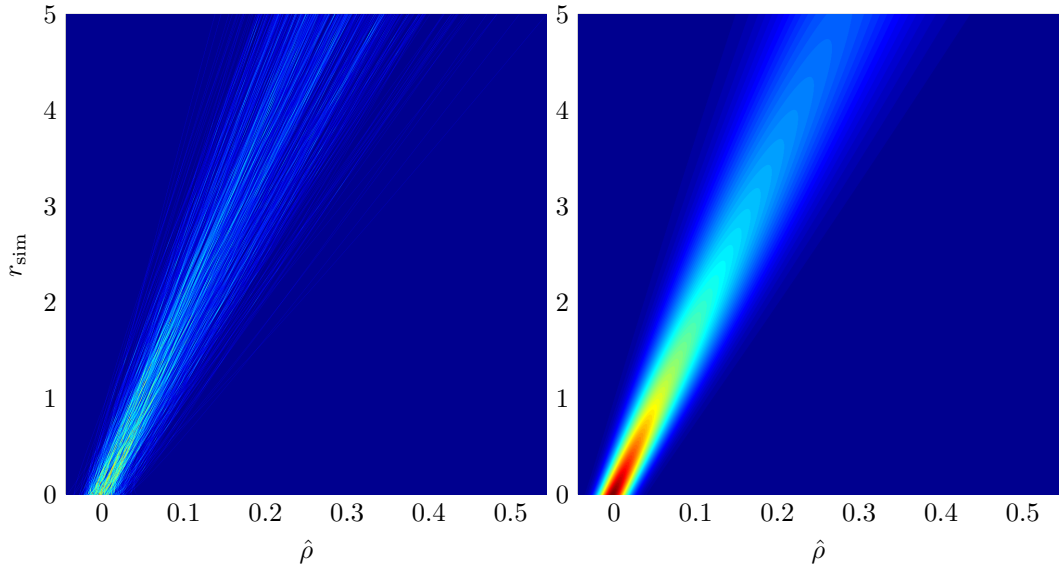


Figure 4.6: Simulated outcomes for BICEP1 three-year signal-plus-noise simulations for a range of input model parameters. *Left:* Histogrammed outcomes of $\hat{\rho}$ for a range of $r_{\rm sim}$ between 0 and 5. We observe that as $r_{\rm sim}$ increases, both the mean and the variance of $\hat{\rho}$ increase. We also observe that outcomes of $\hat{\rho}$ for distinct values of $r_{\rm sim}$ are not statistically independent because of our computational shortcut, resulting in what look like rays in the two-dimensional histogram. *Right:* χ^2 fitted probabilities to outcomes of $\hat{\rho}$ from input values of $r_{\rm sim}$. The distributions are row-normalized such that the probability $P(\hat{\rho}\,|\,r_{\rm sim})$ integrates to unity. Posterior probabilities can be calculated by taking a vertical slice through the distribution (corresponding to one's observed outcome of $\hat{\rho}$) after applying a prior and re-normalizing.

normalized probability of $\hat{\rho}$, given that selection of $r_{\rm sim}$. A vertical slice will yield the likelihood for that selection of $\hat{\rho}$, $\mathcal{L}(r_{\rm sim}\,|\,\hat{\rho})$.

The posterior probability, $P(r\,|\,\hat{\rho})$, is calculated via Bayes' theorem:

$$P(r\,|\,\hat{\rho}) = \frac{\mathcal{L}(r\,|\,\hat{\rho})P(r)}{P(\hat{\rho})}. \tag{4.75}$$

Here, $P(r)$ is the "prior" on r and $P(\hat{\rho})$ is required for normalization. Posterior probabilities can be regarded as evaluations along the vertical axis in Figure 4.6 for a given observation, $\hat{\rho}_i$. In this picture, applying a prior can be thought of as weighting along the vertical axis, and dividing by $P(\hat{\rho})$ amounts to re-normalizing, such that the outcomes integrate to unity in columns of $\hat{\rho}$ rather than rows of $r_{\rm sim}$.

4.7 Systematic uncertainty from absolute calibration and beam window functions

There are two sources of systematic uncertainty that must be included in our probability distribution functions for $\hat{\rho}$: absolute calibration uncertainty and beam uncertainty. These two potential sources of systematic uncertainty deserve a separate treatment because they are multiplicative in the power spectra, unlike other sources of systematics from temperature-to-polarization and E-to-B leakage, which are additive. While potential sources of instrumental polarization can be demonstrated to be negligible relative to the instrument sensitivity, the absolute calibration and beam window function error must be accounted for regardless of the measurement uncertainty in the power spectra.

This description closely follows a treatment developed by Colin Bischoff, which was reported internally in Bischoff 2013. We treat the absolute calibration uncertainty as an ℓ-independent Gaussian error on the amplitude of the power spectrum. If we let $\tilde{g}_{\rm abs}$ and $\tilde{g}_{\rm best}$ be the true absolute calibration and our best estimate of the absolute calibration, respectively, then we consider some error ϵ such that:

$$\tilde{g}_{\rm best} = (1+\epsilon)\tilde{g}_{\rm abs}. \tag{4.76}$$

We assume that ϵ is Gaussian distributed with a mean and variance of zero and σ_g^2, respectively. Since the absolute calibration is simply an overall scaling of the binned power spectrum, the impact on the value of $\hat{\rho}$ will also be an overall scaling of the signal component of $\hat{\rho}$, while the noise component of $\hat{\rho}$ remains unchanged.[10] This is because the noise is scaled from the data directly, so no absolute calibration factor is required to measure the simulated against the real data. This is obviously not the case, however, for the signal component. In the limit that the sample variance is a negligible fraction of the total uncertainty on $\hat{\rho}$, the variance of $\hat{\rho}$ will remain unchanged with changes to the absolute calibration uncertainty.

Since the probability distribution of outcomes $\hat{\rho}$ is well-described by a χ^2 distribution, we can change the effective number of degrees of freedom of the χ^2 fit to account for our additional uncertainty from absolute calibration. The effective number of degrees of freedom $n_{\rm dof}$ of the distribution of $\hat{\rho}$ is calculated as (Bischoff 2013):

$$n_{\rm dof} = \frac{2(\langle\hat{\rho}\rangle + \hat{\rho}_{\rm bias})^2}{\langle(\hat{\rho})^2\rangle - \langle\hat{\rho}\rangle^2}. \tag{4.77}$$

Here, $\hat{\rho}_{\rm bias}$ is the bias on ρ that would be present had we not noise de-biased and E-to-B de-biased at the outset (which is why it is added rather than subtracted[11]), which is unchanged by absolute

[10] As in previous sections, we use hats to distinguish estimates from the underlying "true" value.
[11] This is similar to the old x-offset used in the offset log-normal approximation

calibration error. With the addition of uncertainty in the absolute calibration, the effective number of degrees of freedom becomes:

$$n'_{\rm dof} = \frac{2(\langle\hat{\rho}'\rangle + \hat{\rho}_{\rm bias})^2}{\langle(\hat{\rho}')^2\rangle - \langle\hat{\rho}'\rangle^2}. \tag{4.78}$$

Here, ρ' is the adjusted value of ρ in the presence of non-zero absolute calibration uncertainty.

Considering the effect of the absolute calibration error on $\langle\hat{\rho}'\rangle$ (the expectation value of $\hat{\rho}'$), since we assume the mean value of ϵ is zero, the expectation value of ρ' will similarly remain unchanged ($\langle\hat{\rho}'\rangle = \langle\hat{\rho}\rangle$). The variance, however, will increase as our uncertainty in the absolute calibration grows (as σ_g^2 increases).

We now consider the expectation value of ρ'^2:

$$\langle(\hat{\rho}')^2\rangle = \langle(\hat{\rho})^2\rangle + \sigma_g^2\langle(\hat{\rho}_{\rm sig})^2\rangle. \tag{4.79}$$

Here, $\langle(\hat{\rho}_{\rm sig})^2\rangle$ is the expectation value of the square of the signal-only component of $\hat{\rho}$, $\hat{\rho}_{\rm sig}$. For the case of $r_{\rm sim} = 0$, $\langle(\hat{\rho}_{\rm sig})^2\rangle$ is still non-zero, due to the variance introduced from E-to-B leakage. This allows us to re-write the effective number of degrees of freedom as:

$$n'_{\rm dof} = \frac{2(\langle\hat{\rho}\rangle + \hat{\rho}_{\rm bias})^2}{\langle(\hat{\rho})^2\rangle + \sigma_g^2\langle(\hat{\rho}_{\rm sig})^2\rangle - \langle\hat{\rho}\rangle^2}. \tag{4.80}$$

Each of these components is readily available from signal-plus-noise simulations.

We treat the systematic uncertainty from the beam window function as an ℓ-dependent scaling of the binned bandpowers. The same arguments apply as before: we can account for the additional uncertainty by changing the effective number of degrees of freedom for the χ^2 probability distribution that describes outcomes of $\hat{\rho}$. Also as before, the uncertainty impacts the signal variance contribution to the total uncertainty only. Unlike before, however, the uncertainty is no longer a simple scaling of ρ because of the ℓ dependence of the error. Instead, our estimate $\hat{\rho}$ will be modified by the beam uncertainty according to:

$$\hat{\rho}' = \frac{\langle\mathcal{C}_b^{BB}\rangle^\top\mathcal{N}^{-1}((1+\epsilon_b)\hat{\mathcal{C}}_b^{BB})}{\langle\mathcal{C}_b^{BB}\rangle^\top\mathcal{N}^{-1}\langle\mathcal{C}_b^{BB}\rangle}. \tag{4.81}$$

Here, ϵ_b is the beam error for bin b. As before, we assume ϵ_b has a mean of zero (at each bin) and a variance σ_b, and, since the mean of ϵ_b is assumed to be zero, the expectation value of $\hat{\rho}'$ will remain unchanged. The expectation value $\langle(\hat{\rho}')^2\rangle$ can be calculated in terms of $\hat{\rho}$ and σ_b as (for details, please refer to (Bischoff 2013)):

$$\langle(\hat{\rho}')^2\rangle = \langle(\hat{\rho})^2\rangle + \langle(\Sigma_b\sigma_b w_b\hat{\mathcal{C}}_b^{\rm sig})^2\rangle, \tag{4.82}$$

where $\hat{\mathcal{C}}_b^{\rm sig}$ is a single signal-only realization, and the expectation value is calculated over all realizations. The normalized weight vector w_b is calculated as:

$$w_b = \frac{\langle \mathcal{C}_b^{BB} \rangle^\top \mathcal{N}^{-1}}{\langle \mathcal{C}_b^{BB} \rangle^\top \mathcal{N}^{-1} \langle \mathcal{C}_b^{BB} \rangle}. \tag{4.83}$$

The modified effective number of degrees of freedom then becomes:

$$n'_{\rm dof} = \frac{2(\langle \hat{\rho} \rangle + \hat{\rho}_{\rm bias})^2}{\langle (\hat{\rho})^2 \rangle + \langle (\Sigma_b \sigma_b w_b \hat{\mathcal{C}}_b^{\rm sig})^2 \rangle - \langle \hat{\rho} \rangle^2}. \tag{4.84}$$

4.8 Neyman confidence intervals

From the distribution of outcomes plotted in Figure 4.6, we can explicitly evaluate frequentist acceptance intervals for each discrete value of $r_{\rm sim}$. (Settle down, Bayesians—we will get to you in a moment). The frequentist interval is a simple statistic: At each value of $r_{\rm sim}$, we ask, what band in $\hat{\rho}$ can I define that contains, say, 95% of the most likely outcomes? There are a number of possible criteria that can be used to define such a band—minimum width, one sided, etc. We choose the Feldman-Cousins likelihood ratio, described in Feldman and Cousins 1998, because the interval smoothly transitions between a one-sided and two-sided interval (i.e. between an upper limit and a detection).

The Feldman-Cousins interval evaluates outcomes based on their likelihood ratio $\mathcal{R}$, defined by:

$$\mathcal{R} = \frac{\mathcal{L}(r_{\rm sim} \,|\, \hat{\rho})}{\mathcal{L}(r_{\rm best} \,|\, \hat{\rho})}. \tag{4.85}$$

Here, $r_{\rm best}$ is the value of $r_{\rm sim}$ that maximizes $\mathcal{L}(r_{\rm sim} \,|\, \hat{\rho})$ for a given value of $\hat{\rho}$. Since the likelihood does not extend to negative r, $r_{\rm best}$ likewise cannot be negative. To develop some intuition, let us consider $\mathcal{R}$ in two limiting cases. First, we consider an unlikely highly negative excursion for $r_{\rm sim} = 0$ (an outcome on the far lefthand side of Figure 4.5). The likelihood for that outcome is small, so the numerator will be small. However, for this highly negative point, $r_{\rm best} = r_{\rm sim} = 0$. As a result, the likelihood ratio evaluates to unity. Now we consider a positive excursion (an outcome on the far righthand side of Figure 4.5). The numerator is the same, but now $r_{\rm best}$ is no longer zero, but instead some positive value of r. The likelihood of that outcome will be higher for some larger value of r, and, as a result, $\mathcal{R} < 1$. If we draw an acceptance interval sorted by the likelihood ratio, we see that it establishes a one-sided interval. All negative excursions will be included in the acceptance band (because even though they are negative, they are still most consistent with $r = 0$). Higher excursions will be excluded, thus establishing an upper bound.

If we instead consider some relatively large value of $r_{\rm sim}$, say $r_{\rm sim} = 2$, we can similarly examine the behavior of the likelihood ratio. A negative fluctuation (relative to the mean), which we will call

$\hat{\rho}_{\rm neg}$, will yield a value of $r_{\rm best}$ that is less than $r_{\rm sim}$. Correspondingly, $\mathcal{L}(r_{\rm sim} \,|\, \hat{\rho}_{\rm neg})$ will be less than $\mathcal{L}(r_{\rm best} \,|\, \hat{\rho}_{\rm neg})$ for the simple reason that $\hat{\rho}_{\rm neg}$ favors smaller values of $r_{\rm sim}$. As a result, the likelihood ratio will be < 1. The same argument holds for large fluctuations in $\hat{\rho}$, but instead $r_{\rm best} > r_{\rm sim}$. When we establish an acceptance interval that contains 95% of outcomes sorted by their likelihood ratio, the interval will be two-sided, placing both lower and upper-bounds on the parameter.

Note that the primary reason to use the Feldman-Cousins construction is because of the physical boundary at $r = 0$. Because we define our parameter to be strictly positive, a minimum-width criterion has very little physical meaning. The likelihood ratio is a convenient recipe for ordering outcomes to establish acceptance intervals that transition smoothly from two-sided to one-sided intervals when near a physically motivated boundary for the parameter.

After calculating acceptance intervals for every $r_{\rm sim}$, we evaluate confidence intervals by spanning a range of $r_{\rm sim}$ (amounting to a vertical evaluation in Figure 4.6). For a discrete value of $\hat{\rho}$ (corresponding to an observation), we ask whether the value falls within the acceptance interval for each value of $r_{\rm sim}$ considered. The result is quoted as a 95% confidence interval in r, illustrated in Figure 4.7.

4.9 Bayesian credible intervals

Bayesian credible intervals are calculated by integrating the posterior probability (defined in Equation 4.75). In order to calculate the posterior probability, we have to apply some prior on r, that is, we must plug in some value for $P(r)$ into Equation 4.75. This provides a mechanism whereby independent measurements can be combined to place meaningful constraints on the parameter. For instance, observations of the temperature power spectrum favor $r < 0.11$ (Story et al. 2012), which can be used as a prior. In the BICEP1 three-year analysis, a flat prior on r is assumed. This choice is not motivated by a specific physical model, which may ruffle the feathers of statisticians who insist that priors should be used to convey knowledge of the parameter, rather than ignorance.

These objections notwithstanding, we calculate a posterior probability assuming a flat prior on r. With $P(r \,|\, \hat{\rho})$, we can calculate Bayesian credible intervals by simply integrating the posterior probability up to 95%. We calculate Bayesian credible intervals as strictly upper limits (evaluating the integral positively from $r = 0$).

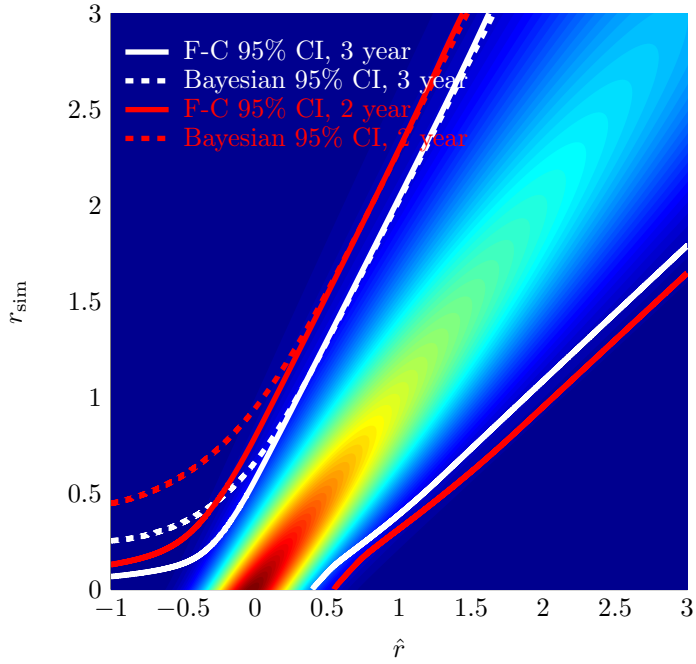


Figure 4.7: Confidence intervals from BICEP1 simulations of the two- and three-year data selections. Bayesian credible intervals are indicated with dashed lines and are strictly upper limits. Feldman-Cousins intervals are indicated with solid lines and transition from one-sided to two-sided intervals. The background color scale are the row-normalized probabilities $P(\hat{r}\,|\,r_{\rm sim})$. These likelihoods include systematic uncertainty from the absolute calibration and beam, as described in Section 4.7. To arrive at parameter constrains from the real data, we draw a vertical line corresponding to the observed value of $\hat{\rho}$ and read off the appropriate Bayesian and frequentist intervals. For negative values of $\hat{r}$, the Bayesian interval is significantly more conservative than the Feldman-Cousins interval.

Chapter 5

Results from BICEP1 and preview of future BICEP2 analyses

In this chapter, we will present new results from the BICEP1 three-year analysis, review the current state of the BICEP2 analysis, and provide a preview of future BICEP2 analyses. The three-year BICEP1 analysis is presented in Barkats et al. 2013, which is, as of this writing, in preparation. We will begin by describing the application of the deprojection algorithm to the BICEP1 three-year data set and present improved constraints on the level of excess BB bandpower from relative gain mismatch and differential pointing. We will also consider the effects of the algorithm on the BB spectra, including the additional loss of information and increased bandpower suppression.

As in the Chiang et al. two-year analysis, the BICEP1 three-year BB bandpowers have been demonstrated to be noise dominated. This is evidenced by the fact that BICEP1 passes a suite of jackknives that are sensitive to various sources of potential systematics, as we will describe. After presenting maps and power spectra from the BICEP1 three-year analysis, we use the analysis methods described in Sections 4.5 through 4.9 to place constraints on the tensor-to-scalar ratio r, the best-ever derived from polarization data alone. We compare these results to Chiang et al. and investigate potential sources of discrepancy.

Finally in this chapter, we will provide a preview of BICEP2 analyses to come. We present three-year maps accompanied by a sensitivity comparison with BICEP1. We also demonstrate the successful application of differential pointing deprojection to the BICEP2 two-year data set—a critical milestone for publication. To conclude, we will outline a path forward for future BICEP2 analyses, highlighting the major outstanding analysis tasks that remain before a BICEP2 initial result is possible. Given the tremendous sensitivity of BICEP2 and the successful strategy pursued by BICEP1, this data set promises exciting new constraints on the B-mode amplitude of the CMB in the near future.

5.1 Overview of analyses

Before describing these analyses in detail, it is useful to first provide some historical context, highlighting how this analysis differs from previous efforts. The primary result from the BICEP1 two-year analysis, described in Chiang et al. 2010 (with the accompanying instrument paper Takahashi et al. 2010), was based on map construction within `Healpix` and relied on power spectrum estimation from the publicly available `Spice` package (Chon et al. 2004). An alternate Matlab-based pipeline was presented in the two-year result paper that was adapted from the QUAD pipeline presented in Pryke et al. 2008. The BICEP1 three-year analysis pipeline, presented in Barkats et al. 2013, incorporates a number of improvements to the Matlab pipeline that will be critical for future analyses of BICEP2 and Keck data. The BICEP1 three-year analysis uses the same low-level reduction as the two-year analysis, but differs in data selection and analysis details.

The most significant differences between the primary two-year Chiang et al. analysis and the three-year Barkats et al. analysis are as follows: in addition to the data selection used in the two-year analysis, the three-year analysis includes not only observations from the 2008 observing season, but also several channel pairs that were excluded from the two-year analysis due to irregular transfer functions. With the additional channels, the three-year analysis includes roughly 50% more integration time. The map construction used for the three-year analysis is a simple flat-sky approximated RA/DEC binned map, whereas the two-year data were binned into `Healpix` maps. Power spectra are estimated from these flat-sky maps using a pseudo-C_ℓ estimator, whereas the two-year analysis was performed on the curved sky using the `Spice` package.

A major development for the BICEP1 three-year analysis is the deprojection analysis described in Section 4.4. As we will describe, this algorithm has been successfully implemented to place much tighter constraints on potential levels of systematic contamination from relative gain mismatch. Additionally, the noise model, used to de-bias the spectra, was improved for the three-year analysis to preserve correlations between azimuthal scans of the telescope. The bandpower window function calculation was also improved to take into account changes in the filtering across bins in ℓ, as described in Section 4.2.4. As we will demonstrate, this particular improvement has a significant impact on error estimates.

The final major improvement for the three-year analysis is the likelihood estimation and parameter constraint method described in Section 4.6. Rather than relying on the offset log-normal likelihood approximation used in the two-year analysis, the three-year analysis arrives at likelihood estimation directly via numerical simulations.

Future BICEP2 and Keck analyses will benefit from the development of these analysis methods. In particular, the deprojection analysis will likely prove critical for removing temperature-to-polarization leakage from differential pointing, and perhaps other sources of leakage as well. Simi-

larly, the improved likelihood and bandpower window function calculation methods will be necessary for accurately reporting constraints and uncertainties.

These are necessary, but not sufficient requirements for an initial BICEP2 analysis. Further developments necessary for an initial BICEP2 result include, but not are limited to: i) improved E/B separation for reduced statistical uncertainty, ii) an expanded deprojection algorithm that includes a wider class of potential sources of systematics, iii) a more detailed foreground analysis, and, iv) potentially, cross-checks against external data sets, including Planck, BICEP1, and Keck.

5.2 Constraints on systematic contamination in the BICEP1 three-year data selection

For many potential sources of systematics, the BICEP1 three-year result relies on the analysis presented in Takahashi et al. 2010. In this instrument characterization paper, it was demonstrated that BB bandpower constraints were limited by statistical uncertainty. Other potential sources of systematic uncertainty contributed spurious polarization at or below $r = 0.1$. For many of these potential sources of systematics, constraints on potential sources of systematics were significantly better than the $r = 0.1$ benchmark. We expect that these spurious polarization limits are equally applicable to the three-year analysis, since the observing pattern and low-level data processing are identical for the additional third year. We refer the reader to Takahashi et al. for a detailed treatment of the potential sources of systematics considered.

The Takahashi et al. analysis included constraints from both instrument-induced and ground-induced systematic contamination. The instrument systematics considered included relative gain mismatch, differential pointing, differential ellipticity, differential beam width, telescope pointing uncertainty, polarization angle uncertainty, far sidelobe pickup, focal plane temperature fluctuations, and optics temperature fluctuations. For the Takahashi et al. analysis, spurious polarization from relative gain mismatch and differential beam properties was estimated by injecting temperature-to-polarization leakage into time streams and carrying through the simulated data to power spectra, as described in Section 4.3.3.

From that analysis, it was demonstrated that only differential pointing and relative gain mismatch could potentially contribute polarization systematics relevant to an $r = 0.1$ benchmark. For this reason, our efforts in the three-year analysis focus on constraining spurious polarization from relative gain mismatch and differential pointing.

5.2.1 Application of the deprojection algorithm to the BICEP1 three-year analysis

The deprojection algorithm described in Section 4.4 has been successfully employed in the BICEP1 three-year analysis to improve constraints on potential systematics from relative gain mismatch and differential pointing. In the two-year analysis, it was noted that uncertainty in the method used to estimate relative gains could potentially contribute systematic contamination at a level comparable to the stated instrument benchmark of $r = 0.1$ (Takahashi et al. 2010).

In the three-year analysis, this is a greater concern in part because of the additional sensitivity, but also because of the addition of the so-called "slow-τ" channels. These detectors, amounting to four channel pairs at 150 GHz and two channel pairs at 100 GHz for the 2007 and 2008 seasons, were excluded because of aberrant detector constants. For this reason, we expect the relative gain mismatch in these channel pairs to potentially be larger than average.

5.2.2 Estimating false polarization from relative gain mismatch prior to template deprojection

The level of potential systematics from relative gain mismatch before and after deprojection is estimated in a series of analysis steps. First, the level of relative gain mismatch is "measured" in the real data by regressing a temperature template from the WMAP seven-year V-band map, as described in Section 4.4.1. A regression coefficient is recovered for each channel pair, and regarded as an estimate of the true relative gain mismatch of the pair. We find that 9 of 20 detector pairs at 150 GHz and 9 of 23 at 100 GHz exceed the benchmark for relative gain mismatch quoted in Takahashi et al. 2010.

These per-channel-pair coefficients are then used as inputs to simulations containing E-mode, but no B-mode polarization. The resulting B-mode polarization in these signal-only simulations contains some combination of E/B mixing from imperfect E/B separation, and spurious B-mode polarization from relative gain mismatch. By generating equivalent simulations containing ideally matched relative gains, it is possible to separate the spurious polarization from relative gain mismatch.

We find that in the three-year data selection, including the "slow-τ" channels, the temperature-to-polarization leakage from relative gain mismatch is predicted to be significant (Figure 5.1). At 150 GHz, excess bandpowers are predicted to exceed the $r = 0.1$ theory spectrum across the ℓ-range of interest. We find that the contamination is predicted to be potentially more significant at 150 GHz than at 100 GHz.

There is a caveat that accompanies this estimate. Because the real data contain noise, the regression coefficients are noisy estimates of the "true" value of the relative gain mismatch. As a

result, there is some uncertainty in the mean level of spurious polarization predicted by this analysis (prior to deprojection). In general, adding noise to the estimate of these coefficients will tend to over-estimate the contamination, since both negative and positive regression coefficients will lead to positive false BB bandpowers. By estimating regression coefficients on noise-only data, we can estimate the amount by which the bandpowers are over-estimated. We can then use these "noise only" coefficients as inputs to forward simulations, from which we can calculate excess BB bandpowers. We find that the excess bandpower estimated from these coefficients is 4% in amplitude, relative to the coefficients estimated from the real data. We therefore conclude that the over-estimation of the false BB bandpowers due to noisy estimates of the regression coefficients is a small effect.

5.2.3 Estimating false B-mode bandpower from relative gain mismatch after template deprojection

Next, we estimate the level of spurious polarization after "deprojecting" relative gain mismatch. In this context, "deprojecting" means subtracting the template scaled by our best estimate of α_{rg} from the time series (details can be found in Section 4.4). To estimate the residual level of false BB bandpowers after deprojection, we follow the same steps as in the previous section. We take the regression coefficients estimated from the real data as inputs to forward simulations. Temperature leakage is injected into the pair-difference time series. After removing the leakage via the deprojection algorithm, we generate BB bandpowers. We then de-bias E/B mixing in the usual manner. We find that the post-deprojection false BB bandpower is over four orders of magnitude lower than the no-deprojection bandpowers (calculated in the previous section) between $37.5 < \ell < 200$ (Figure 5.1). We thus conclude that the deprojection algorithm is extremely efficient at suppressing false B-mode polarization from relative gain mismatch.

We repeat this exercise using simulations containing no E- or B-mode polarization, effectively creating a closed-loop test of the algorithm. In this case, we find that the false polarization from simulated relative gain mismatch is removed perfectly (to within machine precision).

In practice, the dominant residual leakage from relative gain mismatch after deprojection is the result of noise in the external map used to construct the template (in this case, the WMAP seven-year maps). This noise is not accounted for in 5.1 because the deprojection of the simulated data has been performed with a noiseless template. Performing similar simulations in which noise has been injected into the regression template, we find that the suppression of excess bandpower due to spurious polarization is still suppressed by over two orders of magnitude, below equivalent bandpowers for $r = 0.01$ between $30 < \ell < 100$. This residual can be further suppressed in the future by using higher signal-to-noise temperature maps from Planck.

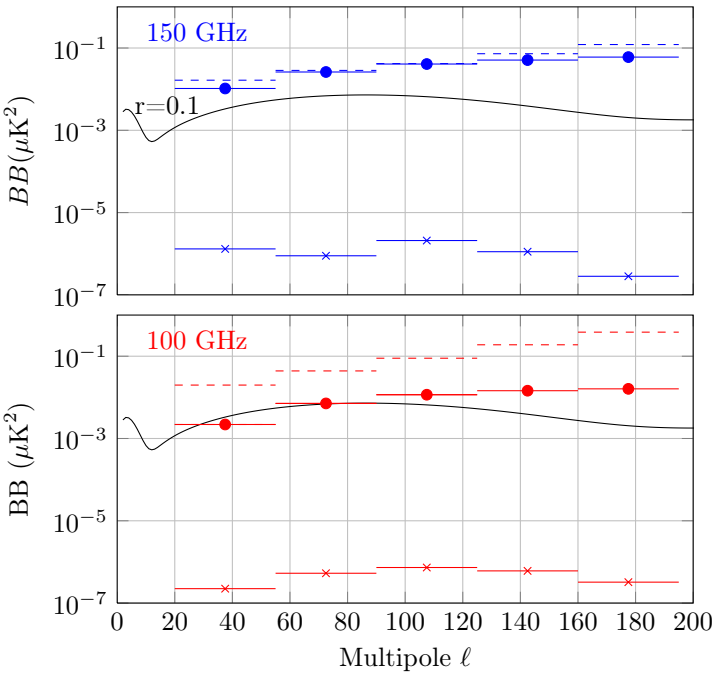


Figure 5.1: The mean relative gain mismatch-induced bandpower, before and after template subtraction, as inferred from signal-only simulations. Dashed lines represent the statistical uncertainty due to noise in the BICEP1 three-year spectra, as inferred from noise-only simulations. Circles represent the mean excess BB bandpower from simulations, assuming relative gain mismatch coefficients inferred from the real data prior to template deprojection. There is a small ($< 5\%$) noise bias on these estimates due to noise in the coefficients, as described in Section 5.2.2. We find that relative gain mismatch contributes significant bandpower relative to $r = 0.1$ (represented by the theory line in black) with no template subtraction. The bandpowers marked with '$\times$' indicate the mean residual BB bandpower after template deprojection. After implementing the algorithm, the false BB contribution in our ℓ-range of interest is suppressed by roughly four orders of magnitude.

5.2.4 False B-mode contamination from differential pointing

We repeat the above exercises, but now considering temperature-to-polarization leakage from differential pointing. Differential pointing templates for x and y displacement are regressed against the real data, spanning all three years of BICEP1 observations. The recovered coefficients are then translated to equivalent pointing centers for A and B and used as inputs to forward simulations. Differential pointing leakage is simulated by simply sampling off of the noiseless signal `Healpix` maps at two different assumed pointing centers for A and B. The resulting pair-difference time series contains the signal from the simulated sky polarization (containing E- but no B-mode power) as well as temperature-to-polarization leakage. The time series is carried forward to power spectrum estimation using the usual steps, including E/B mixing removal and suppression factor correction.

The result is plotted in Figure 5.2. We find the excess bandpower for differential pointing to be much less significant than relative gain mismatch for the bandpowers that contribute the most constraining power on r. At 150 GHz, excess BB power is below the $r = 0.1$ theory curve for the first three bins ($37.5 < \ell < 107.5$), which contain over 95% of BICEP1's constraining power on r. By directly comparing maximum likelihood values of r derived from 100 realizations of signal-only simulations with and without differential pointing leakage, we have confirmed that the resulting potential bias on r is substantially less than $r = 0.1$, and not significant, given BICEP1's statistical uncertainty due to noise. As a result, we find that it is not necessary to actually deproject differential pointing leakage for the BICEP1 three-year analysis.

This may seem in slight tension with the Takahashi et al. analysis, which stated that differential pointing could potentially contribute systematics at a level relevant to an $r = 0.1$ benchmark. However, the benchmark reported in Takahashi et al. is pessimistic; the benchmark was drawn from comparing BB bandpowers directly to an $r = 0.1$ theory curve, and did not account for the relevant weighting between bandpowers. Differential pointing contributes more significantly with increasing ℓ, the same scales at which the bandpower weights are decreasing due to the beam window function.

Although not strictly necessary for the BICEP1 three-year analysis, we demonstrate the ability of the deprojection algorithm to suppress leakage from differential pointing. We proceed in much the same way as before: differential pointing leakage is injected into forward simulations, but is now removed by deprojecting the leakage templates as constructed in Section 4.4.2. The scaled templates are then subtracted from the pair-difference TODs, which are carried forward to power spectra (accounting for E/B mixing and bandpower suppression as before).

The result is plotted in Figure 5.2. The mean residual leakage across 100 realizations is plotted both before and after the differential pointing deprojection has been implemented. We find that the algorithm effectively suppresses simulated leakage by over two orders of magnitude in most ℓ bins, well below the $r = 0.1$ curve.

The algorithm is currently limited by two effects: The first is a non-ideality in the forward

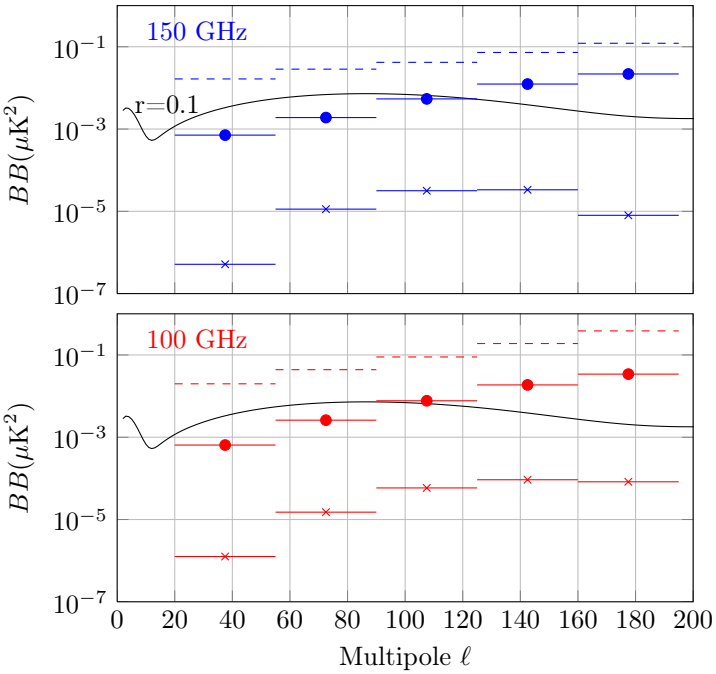


Figure 5.2: The mean excess BB bandpower of signal-only simulations containing differential pointing, both before and after template subtraction. Dashed lines represent the total statistical uncertainty due to noise. The mean excess BB polarization from simulations containing leakage from as-measured differential pointing coefficients before template subtraction is indicated with '$\circ$.' Bandpowers marked with '$\times$' show the mean excess bandpower after template subtraction. We find significant reduction in the potential false BB bandpower. However, owing to the fact that the contamination is primarily at high ℓ, we find that template subtraction is not necessary to achieve the BICEP1 benchmark of $r = 0.1$.

simulation. Detector data streams are simulated by interpolating off of relatively coarse `nside`= 512 `Healpix` maps. The interpolation algorithm suffers boundary-crossing discontinuities at certain locations where the interpolation point switches from one nearest-neighbor pixel to another. When differential pointing is simulated, detector centroids occasionally lie on either side of a pixel boundary. When pair-differenced, the simulated data streams contain a non-physical discontinuity. The second effect limiting the algorithm is noise in the template, as in the case of relative gain mismatch. In this case, however, high-ℓ noise from the beam rolloff is more significant (because the spatial derivative is being calculated). As in the case of relative gain mismatch, this can be improved in future analyses with both higher signal-to-noise and higher resolution measurements.

5.2.5 Bandpower suppression and loss of information

A natural consequence of the deprojection algorithm is increased bandpower variance due to mode removal. The deprojection algorithm effectively discards unique modes in the data corresponding to the leaked modes predicted by the template. As additional degrees of freedom are introduced

(either from introducing additional potential sources of contamination or regressing over shorter timescales), we discard a larger number of modes, thereby increasing the bandpower variance. Here we explore the loss of information from the implementation of this algorithm. Since the differential pointing deprojection was not actually implemented in the BICEP1 three-year analysis, we consider only relative gain mismatch.

Bandpower suppression is assessed by comparing two suites of signal-only simulations, containing cosmological E- but no B-mode power (and neither containing simulated systematics). In the first, bandpowers are estimated in the usual manner, but with no deprojection. In the second, relative gain mismatch deprojection is applied to the pair-difference data. By comparing the ratio of the raw bandpowers, we can assess the fractional increase in the bandpower suppression. We find that the increase is small. At both 100 and 150 GHz, the additional bandpower suppression is less than 4% (Figure 5.3). The bandpower suppression from the deprojection algorithm is subdominant to polynomial filtering and ground subtraction (at low ℓ), and the beam window function (at high ℓ) as can be seen in Figure 4.2.

To assess the additional loss of information, we again generate two suites of simulations, this time containing noise only. One of the noise simulations is subjected to relative gain mismatch deprojection, while the other is not. By comparing the ratio of the standard deviation at each ℓ bin, we find that the additional loss of information is also small. The fractional increase in the error bars is less than 2% across our ℓ range of interest (Figure 5.3).

5.3 Data self-consistency tests for the BICEP1 three-year analysis

Power spectra self-consistency is tested with a series of jackknives, in which data is split according to various data selection criteria and differenced. Spectra constructed from jackknife maps are tested against equivalent signal-plus-noise simulations. Signal that is isolated to one half of the jackknife will result in statistical failures and can be used to identify signal that is not fixed on the sky. The BICEP1 analysis employs a large number of these jackknife tests, each sensitive to different sources of spurious signal. Among BICEP1's arsenal of jackknife tests, the unique telescope design allows for a DK-splitting of the data, in which data between 180 degree rotations are differenced. This is a particularly powerful test, as many classes of instrumental polarization will tend to rotate with DK, whereas the polarization field will rotate with $2DK$.

We will first describe each of the jackknife tests employed by BICEP1, then report results of statistical tests against signal-plus-noise simulations.

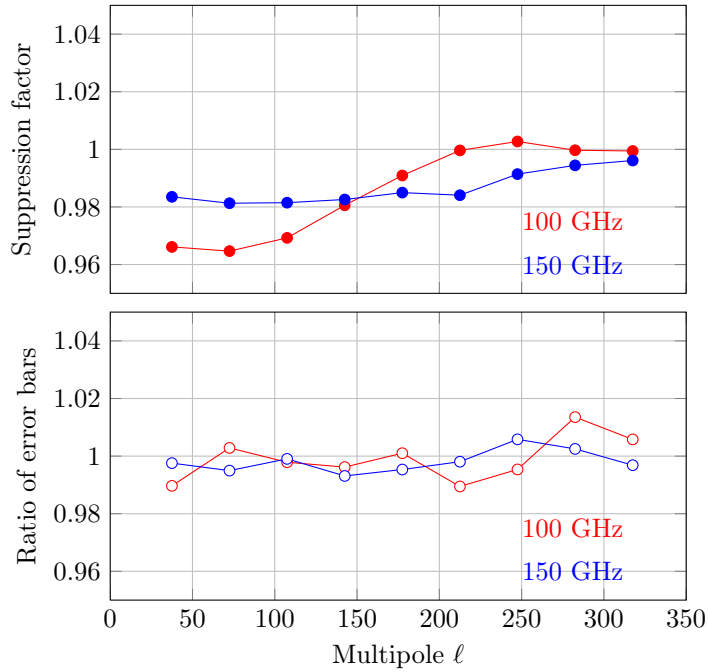


Figure 5.3: Additional bandpower suppression and loss of information from relative gain mismatch deprojection. At both 100 and 150 GHz, the additional bandpower suppression is less than 4%, and subdominant to polynomial filtering, ground subtraction, and beam rolloff. The additional loss of information is similarly small; less than 2% at both 100 and 150GHz. Error bar ratios that exceed 1 are the product of random upward noise fluctuations and the limited number of simulation realizations.

5.3.1 Jackknife tests

The BICEP1 analysis consists of a total of six jackknife tests calculated for each spectra (TT, TE, EE, BB, TB, and EB). Five of these jackknives are calculated separately for 100×100 GHz, 150×150 GHz, and 100×150 GHz. The final frequency jackknife is calculated by splitting the data according to frequency, which results in only a single jackknife for each spectra.

The first jackknife splits the data according to DK angle. Observations separated by 180 degrees are differenced. This jackknife is sensitive to a wide variety of instrument-fixed sources of false polarization, but is particularly sensitive to leakage from differential pointing. The second jackknife splits the data according to scan direction, which serves as a powerful probe of spurious signal from both mismatched detector transfer functions and thermal fluctuations on the focal plane. The third jackknife is a simple temporal split of the data, in which the first half and second half of observations are compared. This jackknife is naturally sensitive to changes to the receiver between seasons, descriptions of which can be found in Chiang et al.

The fourth jackknife is the so-called "focal plane" jackknife. In BICEP1, detectors are aligned to have sensitivity to either Q or U on the sky at $DK = 0$. The focal plane jackknife constructs maps separately for pixels oriented with Q and U on the focal plane, and can serve as a test against instrumental polarization. The fifth jackknife splits the data according to elevation coverage. The BICEP1 scan strategy acquires scansets that alternately cover the upper and lower half of the field. The elevation coverage jackknife differences observations of the upper and lower half of the field and serves as a test of, among other things, ground-fixed signal. The final jackknife test is the aforementioned frequency jackknife, which simply compares observations at 100 and 150 GHz. In addition to testing for instrumental polarization, this can also be used as a check against foregrounds.

5.3.2 Jackknife results

For each of the jackknife tests described above, we compute a χ^2 statistic for the real data and for 499 signal-plus-noise simulations across binned bandpowers (separately for TT, TE, EE, BB, TB, and EB). The χ^2 is calculated for the real jackknifed bandpowers $\hat{\mathcal{C}}_b$, using the mean of the signal-plus-noise simulations $\langle \hat{\mathcal{C}}_b^{\mathrm{s+n}} \rangle$ and the bandpower covariance matrix M:

$$\chi^2 = (\hat{\mathcal{C}}_b - \langle \hat{\mathcal{C}}_b^{\mathrm{s+n}} \rangle)^\top M^{-1} (\hat{\mathcal{C}}_b - \langle \hat{\mathcal{C}}_b^{\mathrm{s+n}} \rangle). \tag{5.1}$$

This χ^2 distribution has 9 degrees of freedom, as we are considering 9 bins. The jackknife bandpowers are assumed to be Gaussian distributed, which is hopefully a safe assumption, since the jackknife maps should be noise-dominated. By computing the same statistic for each simulation realization, we can compute a probability to exceed (PTE) the observed value against the signal-plus-noise simulations. We perform this separately for each spectra (TT, TE, EE, BB, TB, and EB). We say

that we pass a jackknife when the PTE is neither exceedingly high or low. We can additionally test that the PTEs calculated for each of the different frequencies and jackknife tests described above yield a roughly uniform distribution between 0 and 1. The jackknife PTEs for the BICEP1 three-year analysis are given in Table 5.1.

Historically, we find that we routinely fail various jackknife tests involving the temperature spectrum, for both the auto and the cross spectrum. This is not entirely unexpected, given the enormous signal-to-noise in the temperature map. The failures are believed to be caused by imperfections in modeling the noise used in the simulated temperature map. Fractionally small errors in the noise model can lead to imperfectly canceled signal in temperature jackknife maps. This can also cause failures in temperature cross spectra, including TE and TB. This is due to random noise fluctuations correlating with the large signal in the temperature jackknife map, causing a statistical failure when compared against sims.

To test that various TE and TB jackknife failures are due to understood imperfect cancellation in the temperature jackknife maps, rather than unknown failures in the polarization jackknife maps, we construct a series of "half-jacks," jackknives constructed by calculating temperature cross-spectra using the un-jackknifed temperature map against the jackknifed polarization map. We find convincing statistical agreement with simulations, suggesting that TE and TB jackknife failures are due to fractionally small jackknife failures in the temperature map.

5.3.3 Sensitivity to data selection

As described in Section 4.1.2, BICEP1 relies on a single cut criterion to exclude data from the final co-addition. The cut criterion is derived from the standard deviation of relative gain corrections. To verify the cut threshold, we examine how bandpowers fluctuate when the cut threshold is adjusted. Our goal is to demonstrate relative insensitivity to the precise cut threshold, and thus the data selection.

To test the sensitivity of bandpowers to data selection, we generated a series of co-additions of real per-phase data maps with gradually tighter cut thresholds. We calculated four such co-additions, excluding 10% (nominal), 19%, 28%, and 37% of the data volume,[1] which we will call `cut1`, `cut2`, etc. We applied these same cuts to signal-plus-noise simulations for comparison. The simulated maps are calculated as "daughter" co-additions of the same parent set. Each co-addition is thus a subset of the same parent distribution. As a result, the simulated data are largely overlapping between `cut1`, `cut2`, etc.

For each of the progressively stricter cuts, we calculate χ^2 values of the differenced spectra (`cut1`$-$`cut2`, `cut1`$-$`cut3`, etc.) for all six possible permutations. After computing χ^2 values, we compute PTE values against a theoretical χ^2 probability distribution. The theoretical probability

[1] Note that this is not the same as data weight, since we are preferentially excluding low-weight data

Jackknife Spectrum	100-auto	150-auto	100×150	150×100
Scan direction				
EE	0.756	0.124	0.575	
BB	0.244	0.246	0.327	
EB	0.679	0.804	0.148	0.391
Elevation coverage				
EE	0.341	0.471	0.581	
BB	0.106	0.581	0.319	
EB	0.335	0.639	0.273	0.764
Boresight angle				
EE	0.733	0.952	0.192	
BB	0.493	0.257	0.836	
EB	0.489	0.251	0.104	0.026
Season split				
EE	0.495	0.156	0.804	
BB	0.230	0.042	0.525	
EB	0.471	0.421	0.918	0.898
Focal plane QU				
EE	0.986	0.411	0.383	
BB	0.287	0.834	0.451	
EB	0.279	0.244	0.784	0.541

Table 5.1: Results of jackknife tests for the BICEP1 three-year analysis. The probability to exceed the χ^2 values from EE, BB, and EB are quoted for each of the jackknife types described in Section 5.3.1. The PTEs are calculated against 499 signal-plus-noise simulations.

distribution is calculated assuming 9 degrees of freedom (since we are combining 9 bandpowers) and scaled by the variance calculated from the signal-plus-noise simulations.

We find broad agreement with simulations, indicating that by tightening the cut threshold, the bandpowers fluctuate by an amount that is consistent with the changing data volume. The exception is the TT spectrum, for which the PTE is exceedingly low. We attribute this to the high signal-to-noise ratio of the TT spectrum. A small fractional error in the noise model can cause a large fluctuation in the differenced spectra, resulting in an inconsistency with simulations. This is particularly relevant for the low-weight data identified by the cut, which we believe to be least well-represented by the noise model. This is not suggestive, however, of an inconsistency in the data itself.

5.4 Foreground analysis

In this section, we review potential foregrounds in our observing field, specifically foregrounds that may contribute substantial B-mode polarization. This summarizes work presented in Chiang et al. 2010 as well as analysis developed by Jamie Tolan, reported internally in Tolan 2013. There are primarily four possible foreground contaminants: polarized dust emission, galactic synchrotron radiation, polarized radio point sources, and lensing of primordial E-mode polarization. We treat each

of these in turn.

Of these four potential foreground contaminants, the polarized dust emission presents the greatest uncertainty, in part because of a lack of high-fidelity polarimetry at millimeter to sub-millimeter wavelengths. Happily, Planck will soon fill this observational void, offering much more precise polarized dust emission estimates in our observing region. Until the Planck polarization maps are made available, we rely on two dust models: the "FDS model" from Finkbeiner et al. 1999 and the Planck pre-launch sky model (or "PSM12") reported in Delabrouille et al. 2012. The thermal dust intensity used in both of these models is very similar (in fact, PSM12 uses the FDS "model 7" directly), derived from DIRBE, IRAS, and FIRAS data. In the foreground analysis in Chiang et al., a fixed polarized fraction was assigned to the FDS dust model. Polarized dust emission is primarily sourced by dust grains aligning with magnetic field lines, preferentially emitting radiation orthogonal to the field lines. PSM12 uses a more sophisticated polarization analysis that makes use of magnetic field models for the galaxy.

While the polarized dust emission at 150 GHz presents some uncertainty, polarized galactic synchrotron emission can be accurately modeled by scaling lower frequency measurements to 150 GHz. Here again, we make use of PSM12, in which synchrotron-dominated maps taken between 480 MHz and 10 GHz are scaled to higher frequencies, using the component separation analysis reported in Miville-Deschênes et al. 2008. We refer the reader to the documentation of PSM12 for complete details (Delabrouille et al. 2012).

The BB bandpower contamination from polarized synchrotron and dust is constrained by "re-observing" various sky models in our observing region and collapsing the resulting maps into BB auto-spectra. First, a `Healpix` map of the polarized foreground emission model is re-smoothed to the nominal resolution of BICEP. Next, simulated TODs are generated by interpolating off of the `Healpix` maps at locations corresponding to the detector pointing (as described in Section 4.3.1). These TODs are then carried forward through the rest of the analysis steps used for the real data: low-level processing, filtering, map construction, and, finally, power spectra estimation.

For polarized dust, we find that re-observed maps of PSM12 predict higher BB bandpowers than assigning a 5% fixed polarized fraction to the FDS model. However, both models are sub-dominant to equivalent bandpowers for $r = 0.1$ for every bin except $\ell = 37.5$, for which the PSM12 dust prediction slightly exceeds the equivalent bandpower for $r = 0.1$. Due to the assumed spectral index of dust, the contamination at 100 GHz is lower by roughly an order of magnitude. At both 100 and 150 GHz, the predicted polarized dust contribution from either PSM12 or a 5% fixed polarized fraction is small compared to the uncertainty due to noise for the BICEP1 three-year analysis.

We apply the same technique to predict the level of BB bandpower contamination due to galactic synchrotron emission in the field. PSM12 synchrotron maps are re-observed at both 100 and 150 GHz and collapsed into power spectra. We find that at 150 GHz, the predicted synchrotron

contribution is $1-2$ orders of magnitude lower than the predicted contribution from polarized dust. At 100 GHz, synchrotron slightly exceeds the dust contribution, but is small compared to BB bandpowers corresponding to $r = 0.1$.

The third component to consider is foreground emission from polarized radio point sources. Due to BICEP1's broad beam, the instrument is relatively insensitive to point sources in the field. There are, however, a number of radio sources listed in the 4.8 GHz Parkes-MIT-NRAO (PMN) catalog (Wright et al. 1996) in our observing region, as well as sources detected at both 100 and 150 GHz by ACBAR and WMAP (Reichardt et al. 2009; Wright et al. 2009). In Chiang et al., it was demonstrated that masking the brightest 27 ACBAR sources and re-computing angular power spectra negligibly impacted BICEP1's measured BB bandpowers.

The final potential foreground contaminant to consider comes from lensed E-modes. Polarization that emerges as pure-E from recombination is mixed into B by gravitational lensing of intervening matter between the observer and the surface of last scattering. The theory spectrum used for the gravitational lensing model is taken from `CAMB`, based on the computational method reported in Challinor and Lewis 2005. From this lensing theory spectrum, we can calculate bandpower expectation values. Unsurprisingly, the lensing bandpower expectation values contribute negligibly to BICEP1's three-year bandpower estimates. To calculate the contribution to BICEP1's estimate of r, we combine the lensing bandpower expectation values using the weights in Table 4.4. This yields a lensing bias on r_0 of $r_{\rm lens} = 0.03$.

5.5 BICEP1 three-year maps and power spectra

In this section, we present maps and power spectra for the BICEP1 three-year data selection. Maps are constructed as described in Section 4.1.3. Map pixels are 0.25° on a side, plotted in a Mollweide projection. For comparison, the frequency co-added map is plotted against the frequency jackknife map. Both maps have been re-smoothed to account for the differing beam window functions. It is apparent from a visual inspection of these maps that both temperature and E-mode polarization are measured with high significance, whereas the B-mode map appears similar between the frequency co-addition and frequency jackknife. To emphasize the E-mode signature, the polarization maps have been apodized with the weight mask to give higher significance to the central region of the map, which has deeper integration. Additionally, the maps have been spectrally filtered with a boxcar passband of $37.5 < \ell < 200$. After excluding calibrations, cryogenic operations, and telescope turnarounds, the weather cut excludes 10% of the remaining data volume. Altogether, we find a 50% increase in the total data volume over the Chiang et al. data release.

We estimate the map depth using noise-only simulations spanning the same data selection as the real data. To estimate the map depth, we calculate the median per-pixel variance for various

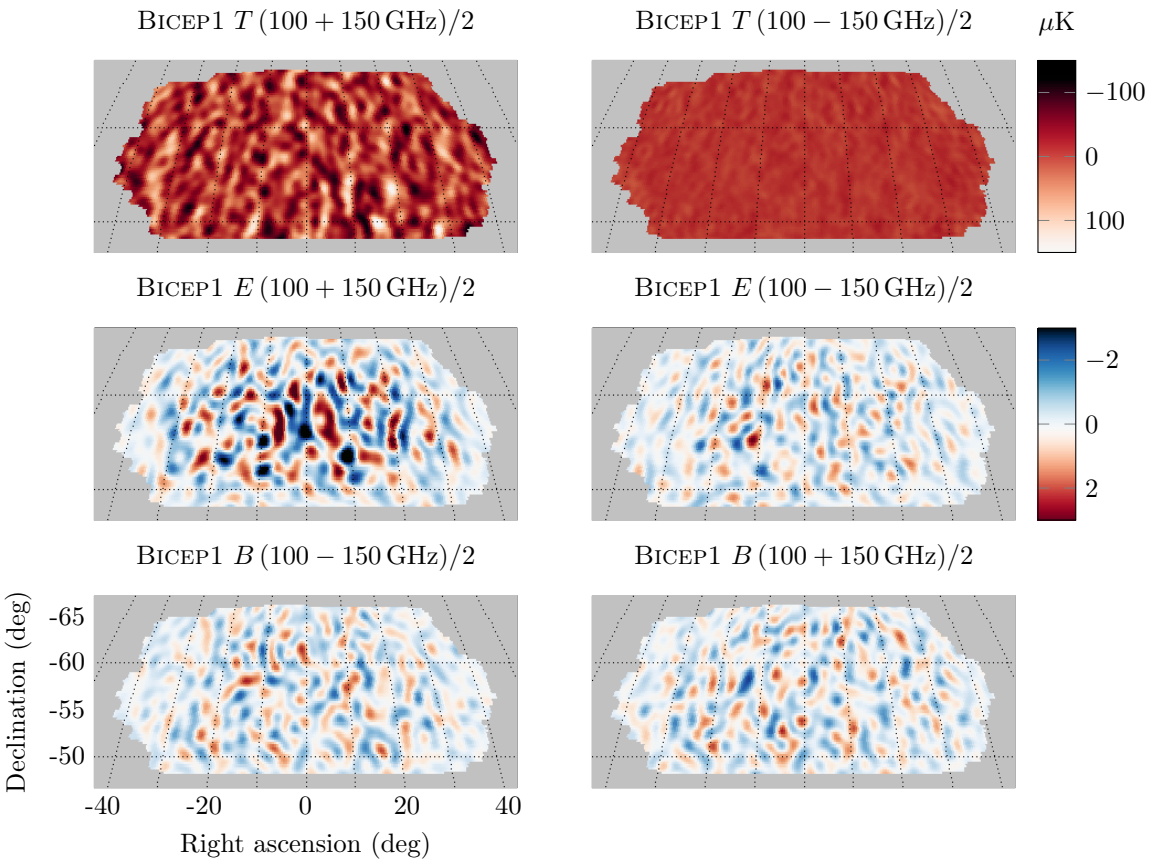


Figure 5.4: Temperature, E-mode, and B-mode maps from three years of BICEP1 observations. The 150 GHz maps have been beam-corrected and re-smoothed to an equivalent 100 GHz beam width. Frequency combined and jackknife maps appear in the left and right columns, respectively. The temperature maps are boxcar-filtered with a passband of $37 < \ell < 280$, while the polarization maps are filtered with a passband of $37 < \ell < 200$.

map-space integration time cuts (the central portion of the map having the highest integration time). We find that in the central region of the Q and U maps, for which there is over 4×10^4 s of per-pixel integration time, the rms noise per square degree is 0.66 μK and 0.50 μK at 100 and 150 GHz, respectively. This central region constitutes 245.6 and 222.7 deg^2 at 100 and 150 GHz, respectively. These noise estimates are roughly consistent with the Chiang et al. noise estimates of 0.81 μK and 0.64 μK (per square degree), given the 50% increase in data volume. This is not an exact one-to-one comparison because the map areas over which the noise is estimated are slightly different. However, this is a small correction. Extending the map area to 556.2 and 415.2 deg^2 (at 100 and 150 GHz), including pixels with over 2×10^4 s of per-pixel integration time, we find a noise per square degree estimate of 0.77 μK and 0.61 μK for 100 and 150 GHz.

Frequency-combined power spectra for the BICEP1 three-year analysis are plotted in Figure 5.5 for T, E, and B auto- and cross-spectra. Theory curves, assuming ΛCDM parameters derived from

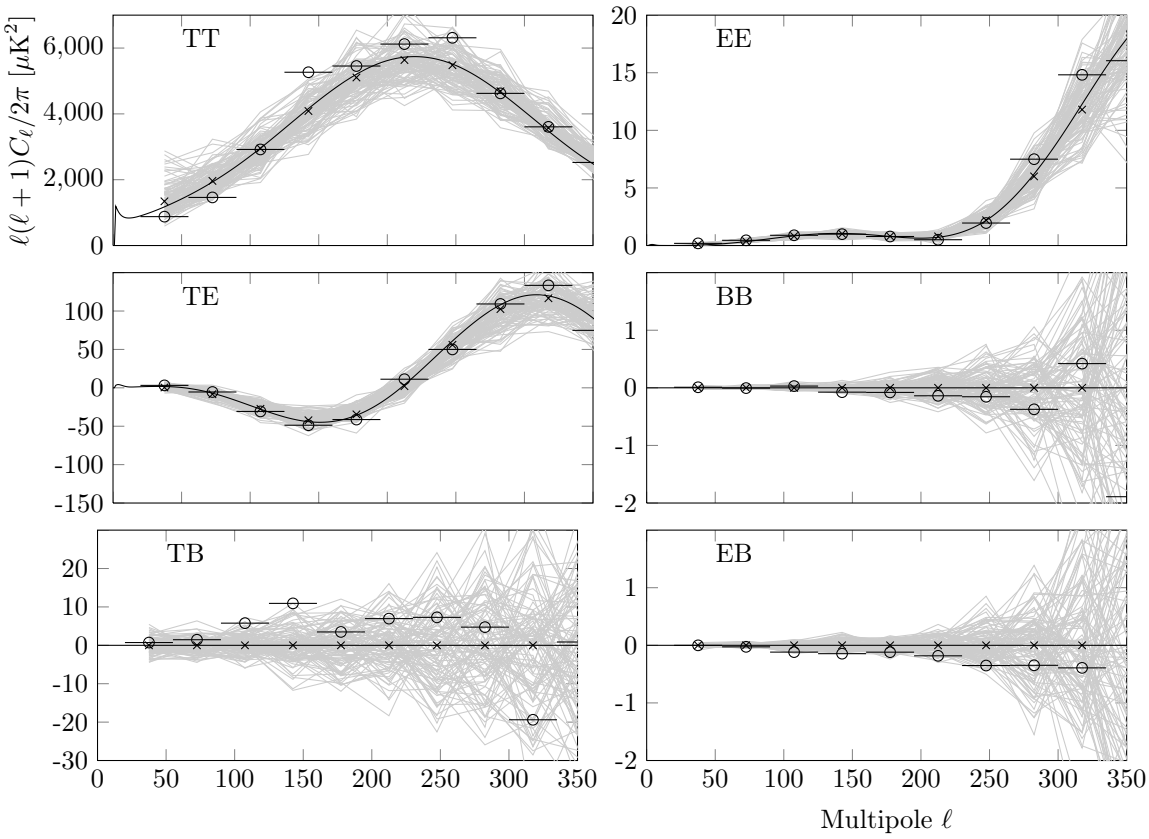


Figure 5.5: Frequency-combined bandpowers for three years of BICEP1 observations. Real data points are indicated with open circles (with horizontal bars indicating bin width), plotted against theory spectrum (solid line) and expectation values (×) for ΛCDM parameters derived from WMAP five-year observations. The gray lines indicate bandpower outcomes for the first 100 signal-plus-noise simulations generated from the same ΛCDM parameters.

WMAP five-year observations, are plotted in black. Binned theory expectation values, as calculated in Equation 4.11, are indicated with '×' markers. The welter of gray lines indicate bandpower outcomes for the first 100 signal-plus-noise simulations, de-biased and suppression-factor corrected using the procedure described in Section 4.2. Real data points are indicated with open circles, with horizontal lines indicating the bin width. A statistical comparison against ΛCDM can be found in Barkats et al. 2013.

In Figure 5.6, we highlight the low-ℓ measurements of the EE and BB frequency-combined spectra. The EE spectra is measured with high significance and is consistent with the simulated distribution calculated from best-fit ΛCDM parameters from the WMAP five-year analysis. The theory spectrum plotted for BICEP1 is for $r = 0$ (no B-mode power). The BICEP1 BB bandpowers appear consistent with zero. Again, a detailed study of statistical consistency with ΛCDM is presented in (Barkats et al. 2013).

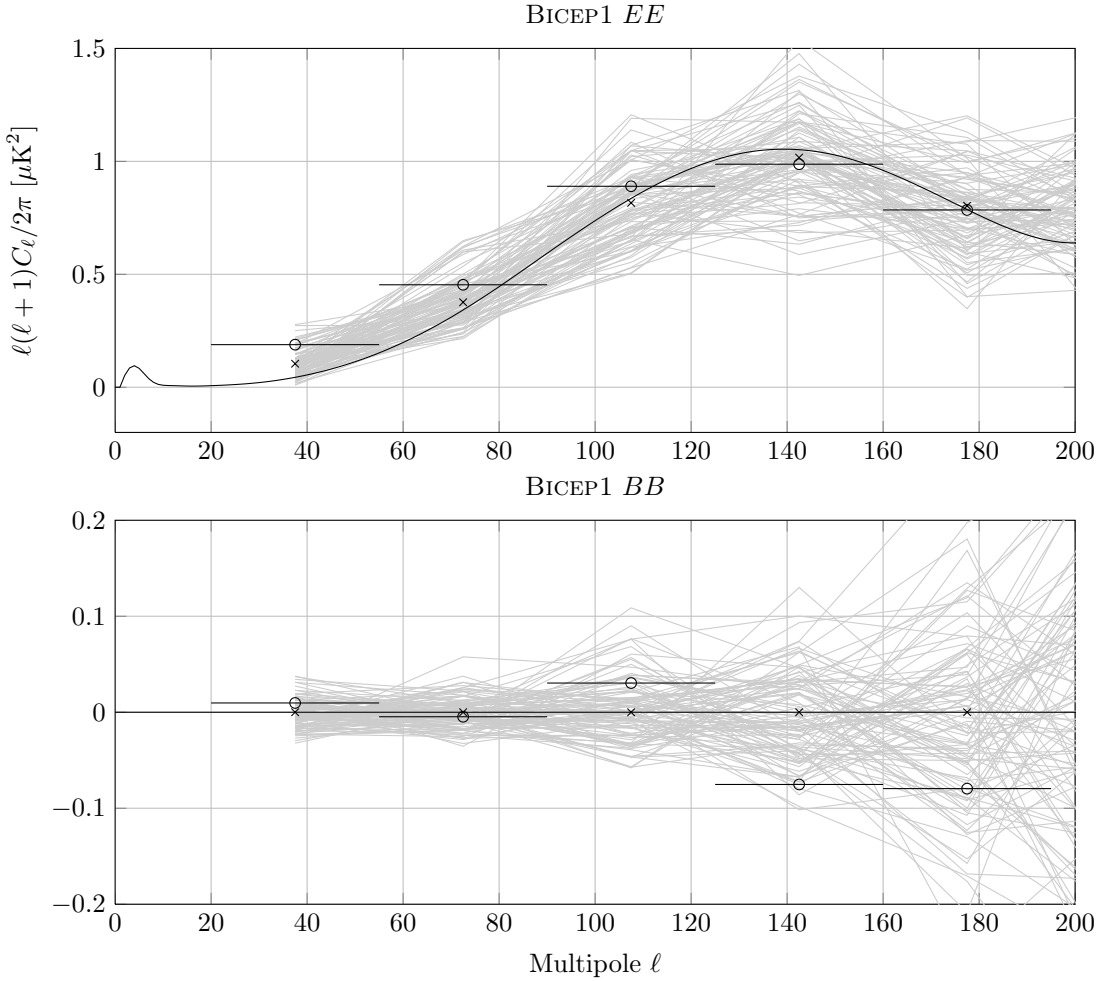


Figure 5.6: Frequency-combined EE and BB spectra from BICEP1 three-year analysis. The same data is plotted as in Figure 5.5, but now with increased scale to illustrate the low-ℓ measurements of EE and BB. The primordial EE spectrum is measured with high significance, whereas the BB spectrum is consistent with zero.

5.6 Constraints on r from the BICEP1 three-year data selection

With the frequency-combined BB bandpowers from Figure 5.5 in hand, we place direct constraints on r, the tensor-to-scalar ratio. Given the large number of analysis changes between the three-year analysis and two-year analysis (the Chiang et al. result), we first repeat the analysis for the two-year data selection to compare constraints on r.

5.6.1 Repeating constraint calculations from the BICEP1 two-year data selection

Using precisely the same data selection as in Chiang et al. 2010, we re-calculate constraints on r, incorporating all of the analysis changes that accompany the three-year analysis. This includes the new noise model, the relative gain mismatch deprojection algorithm, the new bandpower window function estimation method, and the new likelihood estimation method. We begin by calculating frequency-combined bandpowers. Next, we collapse over bins 2-10, taking a weighted linear combination of binned bandpowers as in Equation 4.74, using the weights given in Table 4.4.

This observed weighted linear combination of $\hat{\rho}_0$ is then used to select a posterior likelihood $P(r \,|\, \hat{\rho}_0)$ out of the large family of likelihoods generated from signal-plus-noise simulations of the two-year data selection (including systematic uncertainties in the absolute calibration and beam).

We find that the two-year data selection yields a Bayesian 95% credible interval of $r < 0.80$. Using the construction described in Section 4.5, we calculate a quadratic estimate of r from the two-year BB bandpowers, yielding $r_0 = -0.14 \pm 0.32$. The quoted uncertainty is the square root of the variance of the probability distribution of outcomes in $\hat{\rho}$ for $r_{\rm sim} = 0$. Note that the quadratic estimate of r is not equal to the maximum likelihood of r, which is strictly positive.

The re-analysis of the two-year data selection is not in perfect agreement with Chiang et al.Specifically, the 95% Bayesian credible interval is slightly less constraining ($r < 0.8$ versus $r < 0.72$ in the Chiang et al. analysis). We attribute this difference primarily to improvements in the calculation of the bandpower window function, summarized in Section 4.2.4. In the previous analysis, ignoring the effects of filtering in the bandpower window function caused an overestimation of the bandpower suppression factor, which, in turn, led to an underestimation of the error bar on the first science bin (centered at $\ell = 37.5$).

The best estimate of r quoted in Chiang et al. is $r = 0.02^{+0.31}_{-0.26}$. This is slightly higher than the result from the quadratic estimate of r, but well within quoted uncertainties. We understand the primary differences to be a slightly larger inferred noise bias, given the improved noise model, and differences between the offset log-normal likelihood approximation and the direct quadratic estimator methods.

5.6.2 Improved constraints on r from three years of BICEP1 data

We now extend the data selection to include the full three-year data set, including the "slow-τ" channels. We calculate $\hat{\rho}_0$ using the bandpowers plotted in Figure 5.5. We use this value of $\hat{\rho}$ to construct a posterior likelihood, assuming a flat prior on r. This can be visualized as taking a vertical slice in Figure 5.7 through the family of probabilities constructed for all values of r considered.

As before, by finding the linear scaling between $r_{\rm sim}$ and the mean value of $\hat{\rho}$, we convert $\hat{\rho}_0$ to a quadratic estimate of r. We find that this estimate yields $r_0 = 0.07 \pm 0.23$. Here again, the quoted uncertainty is the square root of the variance of the $\hat{\rho}$ distribution for $r_{\rm sim} = 0$. If we are to regard this as a best estimate of the primordial tensor-to-scalar ratio, we must account for the lensing contribution to the B-mode spectrum. We compute the bias on r due to lensing to be $r_{\rm lens} = 0.03$, as described in Section 5.4. Removing this bias, we find a best-estimate of the primordial tensor-to-scalar ratio to be $r_{\rm prim} = 0.04 \pm 0.23$. This is plotted as a dashed vertical line in Figure 5.7. This observed best-estimate yields a Feldman-Cousins 95% confidence interval of $r < 0.63$ and a Bayesian 95% credible interval of $r < 0.70$.

It might appear suspect that adding 50% more data results in such a marginal improvement in 95% CI on r, but such is the nature of statistics. In fact, comparing with simulations, we find that 15% of the time, the additional data volume will actually lead to a weaker constraint than the two-year data selection, even though the data sets are largely correlated. We therefore conclude that the fact that the three-year analysis yields only a slightly tighter constraint than the two-year analysis is completely consistent with simulations.

5.7 Looking forward: Future BICEP2 analyses

BICEP2 has now concluded three years of observations at the South Pole. This data set contains over 14,000 hours of integration on our primary CMB field. Together with BICEP2's increased mapping speed (over a factor of 10 relative to BICEP1), these observations constitute one of the deepest measurements of the degree-scale polarization of the CMB to date. This unprecedented sensitivity requires unprecedented control of instrumental polarization, ground pickup, sidelobes, magnetic response, and thermal pickup. We seek to demonstrate that each of these potential sources of systematics is well-below statistical uncertainty, and these efforts are ongoing.

The largest potential systematic for BICEP2, mismatch of the pointing centers of detectors within a polarization pair, has been successfully mitigated using the deprojection analysis, as we will describe in this section. A number of other analysis improvements used in the BICEP1 three-year analysis are readily applicable to the BICEP2 pipeline, including the likelihood estimation technique, the improved bandpower window function calculation, and the deprojection of higher-order modes.

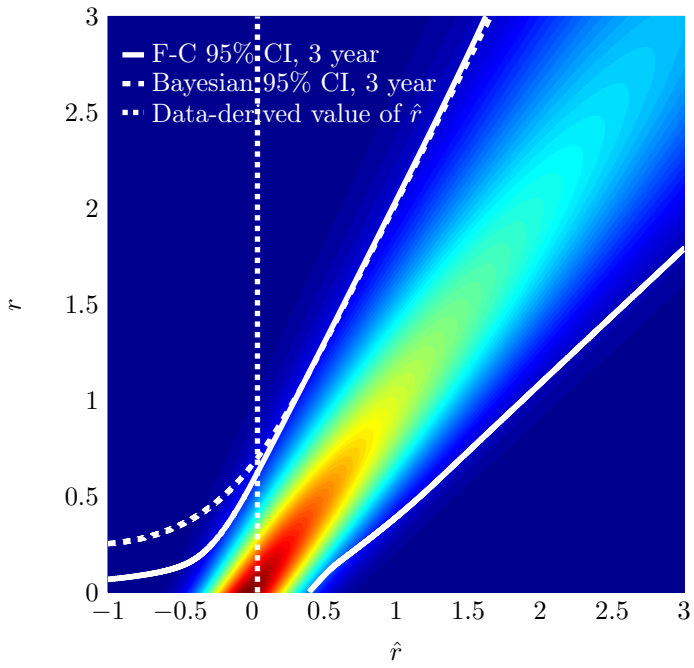


Figure 5.7: Bayesian and frequentist constraints on r from the BICEP1 three-year data analysis. The frequentist confidence band is indicated by the white solid lines, while the 95% Bayesian credible interval is indicated with the dashed line. The observed value of $\hat{r} = 0.04$ is indicated by the dotted vertical line. Bayesian and Feldman-Cousins 95% confidence intervals, given the observed estimate, are $r < 0.7$ and $r < 0.63$, respectively.

Efforts to publish BICEP2 spectra and new constraints on the B-mode polarization amplitude of the CMB are ongoing. While a large number of potential sources of systematics have been excluded through calibration measurements, data simulations, and data jackknives, new classes of potential systematics must be considered. Such possible candidates for potential systematics include, but are not limited to, contamination from cross-talk from the multiplexing architecture, near sidelobes, and magnetic pickup. Many of these sources have been demonstrated to be subdominant, but further analysis work is necessary.

In this section, we will present BICEP2 sensitivity and map depth estimates, discuss the successful application of the deprojection algorithm to BICEP2, present some early analysis E and B-mode maps, and, lastly, discuss the path forward for the BICEP2 analysis.

5.7.1 BICEP2 sensitivity and map depth

Since deployment, BICEP2 has achieved over an order-of-magnitude increase in mapping speed compared to BICEP1. This is owing primarily to three factors: BICEP2's increased detector count, increased per-detector sensitivity (described in detail in Brevik et al. 2010), and increased observing efficiency. By reducing time spent on cryogenic operations and telescope turnarounds, BICEP2 has been able to increase the fractional on-source time relative to BICEP1, as described in detail in Ogburn et al. 2010. Additionally, because of a string of highly successful engineering runs and a very efficient deployment, BICEP2 suffered minimal downtime during the first few months of science data acquisition.

To compare the ultimate achievable sensitivities of BICEP1 and BICEP2, we calculate the average noise in the deepest regions of the Q and U maps. A single map depth statistic is not sufficient for estimating ultimate constraints on r, but rather serves as a point of comparison between the BICEP1 and BICEP2 experiments. For BICEP1, the map depth is calculated for the deepest 200 deg^2 of the field (based on integration time) co-added over three years of observations. Map-based noise estimates are drawn from noise simulations. We find the average Q/U map pixel noise in the deepest 200 deg^2 to be 0.678 μK/deg^2 and 0.50 μK/deg^2 at 100 and 150 GHz, respectively. This is consistent with the 50% increase in data volume over the BICEP1 two-year analysis presented in Chiang et al. 2010, which reported a map depth in Q/U of 0.81 and 0.645 μK/deg^2 at 100 and 150 GHz, respectively, for the same map area.

The BICEP2 three-year Q and U map depth is estimated in the same manner: A map area of 200 deg^2 is selected based on the total integration time (exceeding 6.1×10^5 s). Rather than estimating noise from simulations, we estimate the noise directly from scan-direction Q and U jackknife maps (constructed from differencing left-going and right-going scans). The two map-based noise estimation methods (using scan direction versus noise simulations) will have strong agreement, so long as there is little correlated noise in the left and right going scans. For temperature maps,

this is generally not the case, but for polarization maps, we find this to be true (particularly after ground subtraction).

From the jackknife maps, we find an average Q/U map pixel rms of 0.089 μK/deg^2 in the deepest 200 deg^2 of our map. This is roughly consistent with noise estimates derived from the average per-detector sensitivity quoted in Ogburn et al. 2012 of 15.9 μK $\sqrt{s}$, from which we calculate a Q/U map depth of 0.076 μK/deg^2.

5.7.2 Applying deprojection to BICEP2

As described in Section 3.1.4, BICEP2 suffers from mismatch of detector centroids within a polarized pair (which we call differential pointing). Relative to other types of beam mismatch, differential pointing is by far the largest amplitude. As can be seen from Figure 5.2, the direction and magnitude of the differential pointing has a common-mode component shared between most detectors on the focal plane. For this reason, DK rotation provides a large cancelation of the contamination (because the beam displacement is rotated with respect to the field). Similarly, the DK jackknife, constructed by differencing observations separated by 180 degrees, is a powerful probe of common-mode differential pointing.

The deprojection algorithm, described in Section 4.4.2, has been successfully implemented to suppress temperature-to-polarization leakage from differential pointing. The algorithm has undergone several improvements for the application to the BICEP2 data set. The largest of these improvements is to use spatial derivatives of the Planck 143 GHz temperature maps for the differential pointing templates. The Planck templates benefit from not only higher sensitivity detectors, but also higher resolution relative to the WMAP V-band maps used for the BICEP1 deprojection. Also, higher `nside` maps were used to take full advantage of Planck's higher resolution.

Figure 5.8 illustrates the application of the differential pointing deprojection algorithm to BICEP2 two-year E-mode maps. We examine the DK jackknife map, which, as previously mentioned, has the effect of amplifying the leakage of differential pointing relative to the nominal map. Before applying the deprojection algorithm, the leakage in the DK jackknife is booming – roughly the same amplitude as the E-mode signature itself. We again emphasize that the leakage in the DK jackknife map is enormously amplified relative to the leakage in the non-jackknife map because of the large leakage cancellation that results from DK rotation.

After implementing the deprojection algorithm, we find that the signal in the DK jackknife map is enormously suppressed. We find that the post-deprojection DK jackknife map has comparable amplitude to signal-plus-noise simulations constructed assuming ideally matched pointing centers. Given the additional suppression of the leakage from DK rotation in non-jackknife maps, we find that differential pointing contributes negligible spurious polarization after deprojection.

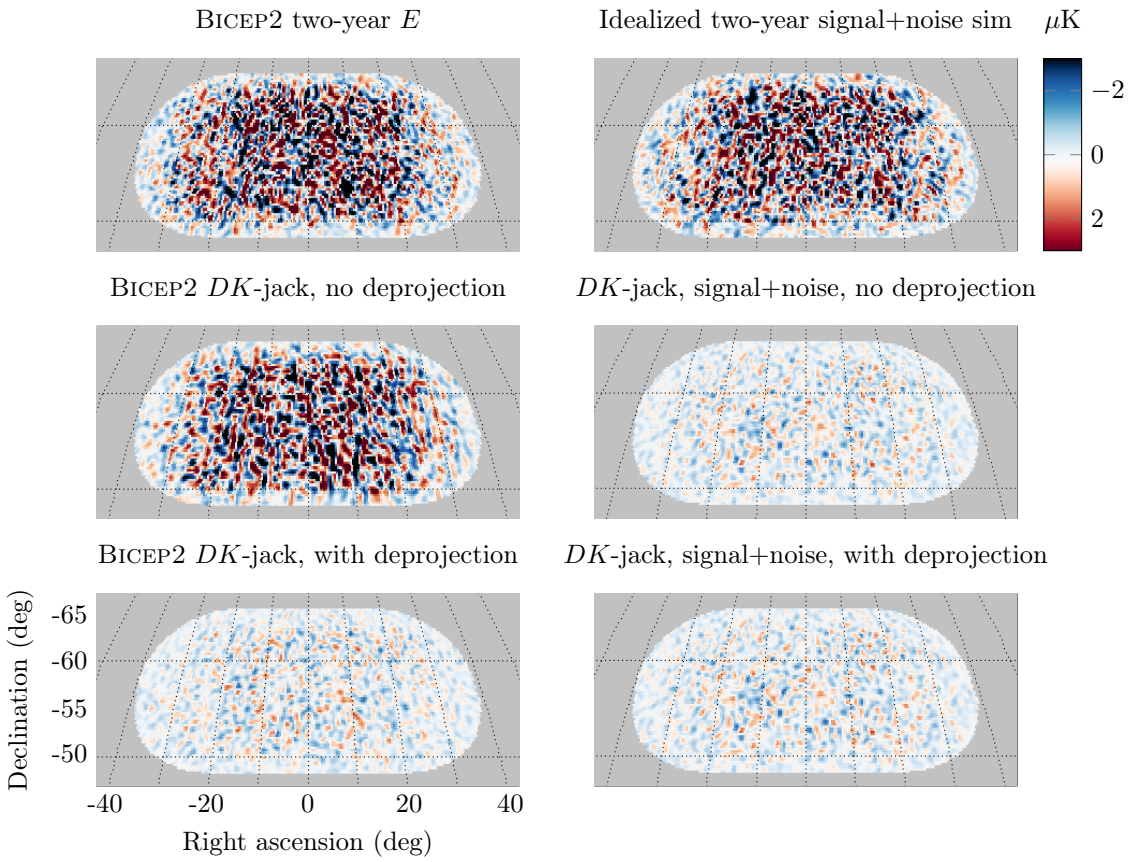


Figure 5.8: BICEP2 two-year E-mode maps, filtered with an ℓ-space spectral boxcar filter from $30 < \ell < 400$. *Top row:* The left-hand map shows the BICEP2 two-year E-mode map (with deprojection). The right-hand map shows a signal-plus-noise simulation, assuming ideally matched detector centroids. *Middle row:* The left-hand map shows the BICEP2 E-mode jackknife map, constructed by differencing observations separated by 180 deg in DK angle. This jackknife has the effect of amplifying common-mode differential pointing leakage. The right-hand map shows the same jackknife for the signal-plus-noise simulation. The signal component cancels exactly, since ideally matched pointing centroids have been assumed. *Bottom row:* The left-hand map shows the same DK jackknife map, but now after implementing the deprojection algorithm for differential pointing using the Planck 143 GHz temperature map. We find that the map amplitude is significantly reduced when deprojection is implemented. The right-hand figure shows the idealized signal-plus-noise simulation with deprojection, which is largely unchanged after deprojection. Given the relative map amplitudes of the idealized signal-plus-noise simulated DK jackknife and the BICEP2 DK jackknife after deprojection, we conclude that the cleaning algorithm is extremely effective at removing spurious polarization.

5.7.3 BICEP2 three-year map comparison

While the final BICEP2 analysis is ongoing, it is illustrative to compare BICEP1 and BICEP2 three-year maps. The BICEP1 maps used for this comparison are combined over 100 and 150 GHz, using the same methods and data selection as described in Section 5.5. The BICEP2 maps are co-added over three years, using the data selection summarized in Tables 4.1 and 4.2, extended to three years

of observations.

As in the case of BICEP1, the BICEP2 E- and B-mode maps have been constructed using a simple Fourier pseudo-$\mathcal{C}_\ell$ estimator on the flat sky, and, as a result, both the BICEP1 and BICEP2 maps contain some non-zero E/B mixing. A comparison of BICEP1 and BICEP2 T, E, and B maps is plotted in Figure 5.9. Both maps have been masked with an identical integration time cut. Owing to BICEP2's higher detector count and higher observing efficiency, the map area observed with equal integration time is somewhat larger for BICEP2. The BICEP1 and BICEP2 maps are not identically filtered; the BICEP1 maps are smoothed to a 0.93° FWHM beam, while the BICEP2 average beam width is 0.52° FWHM. This difference in beam-smoothing is clear in the T maps, but less obvious in the E and B maps, since they have been further filtered with an ℓ-space boxcar filter from $30 < \ell < 200$. The polarization maps have also been apodized toward the map edges, where the map depth degrades.

Both T and E are measured with high significance. Additionally, the BICEP2 B maps are significantly deeper than the equivalent BICEP1 map. In the very near future, these BICEP2 maps promise the most sensitive measurement of the degree-scale polarization of the CMB to date.

To illustrate the physics of the E-mode polarization, we can plot stacked hot and cold temperature anisotropies overlaid with polarization vectors extracted from BICEP2 three-year Q and U maps (Figure 5.10). We first locate local extrema in the map, which has been pre-smoothed to a beam width of $\sigma = 0.5$ degrees. We then stack and average both the temperature and polarization (Q and U) maps over these local maxima and minima, representing the polarization orientation and magnitude with headless vectors. We see a high degree of correlation between the magnitude and orientation of the polarization vectors relative to the hot and cold spots.

We can similarly co-add over hot and cold spots in the Q and U maps. As before, local extrema are identified in the map, but now in Q and U rather than T. Different regions of the temperature and polarization maps are then co-located according to the locations of the extrema and then stacked and averaged. The result is plotted in Figure 5.11. As in Figure 5.10, the background image in each of the four tiles shows the stacked temperature anisotropy. Polarization vectors are overlaid in blue. The physical interpretation of this result is a bit more intuitive: We can immediately see that quadrupolar anisotropies in temperature give rise to linear polarization. The orientation of the polarization vector is directly tied to the orientation of the quadrupole. The temperature anisotropy in $\pm Q$ is somewhat distorted by filtering along the scan direction, which is along the horizontal axis.

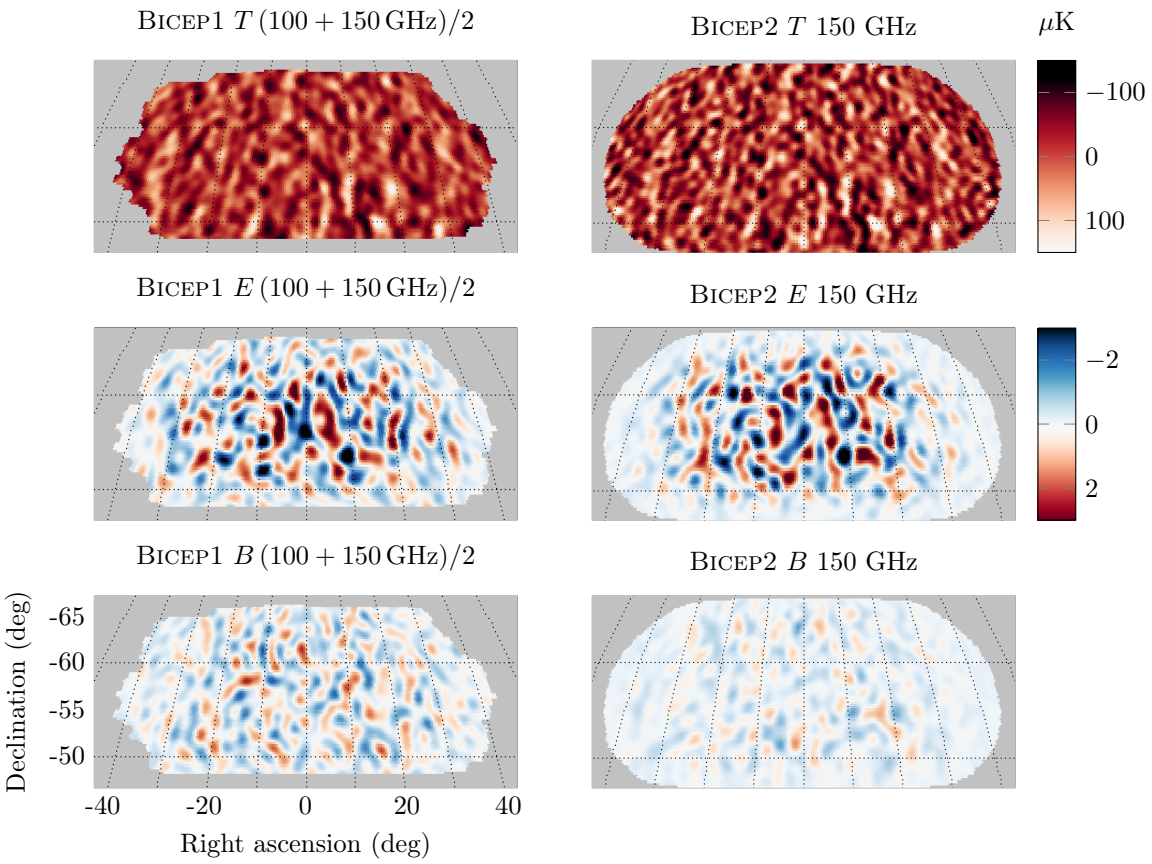


Figure 5.9: T, E and B-mode three-year maps for BICEP1 and BICEP2. The ℓ-space filtering of the maps is not identical between the two sets of maps: The BICEP1 maps have been smoothed to the BICEP1 100 GHz equivalent beam width (0.93°) for frequency combination, while the BICEP2 average beam width is 0.54 degrees. The polarization maps have been further filtered with a boxcar filter from $30 < \ell < 200$.

5.8 Path forward

The BICEP2 analysis is now well underway, with a number of critical milestones already achieved. There are, however, a number of analysis tasks that remain before spectra and parameter constraints can be published. We briefly summarize the path forward for publishing systematics-free BICEP2 spectra.

While, in this thesis, we have measured a number of instrument parameters to high-precision, work remains to translate these parameters to potential systematic biases on r. In the BICEP1 analysis, we took various instrument parameters as inputs to forward simulations, which were then carried forward to power spectra. A similar analysis is necessary for BICEP2. Furthermore, using the ρ construction described in Section 4.5, we can collapse simulated excess bandpowers across ℓ to generate upper limits on the potential level of systematics from relative gain mismatch, differential pointing, differential beam width, and differential ellipticity, both before and after implementing the

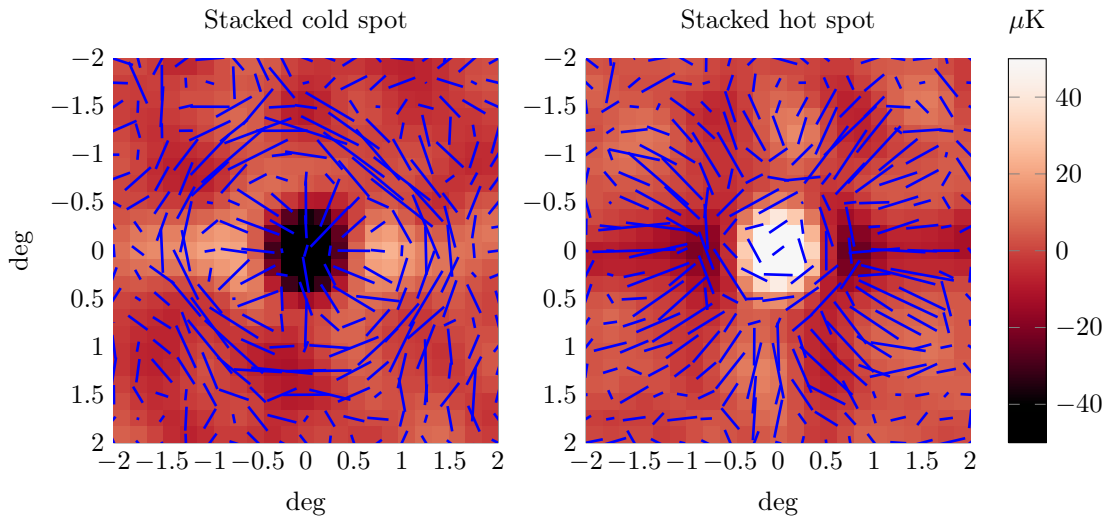


Figure 5.10: BICEP2 three-year stacked hot and cold temperature anisotropies with polarization magnitude and orientation overlaid. TE correlation is demonstrated with high significance. The "curl-less" nature of E-modes is likewise demonstrated.

deprojection algorithm. Similar simulation analyses can be performed to diagnose the contamination due to thermal and magnetic pickup. In all of the cases above, our expectation is that these will be sub-dominant, but a more quantitative analysis is necessary.

While a large number of potential sources of systematics have already been considered and constrained for BICEP2, other sources either loosely constrained or ignored in this thesis may need to be more carefully examined for future analyses. In particular, un-modeled residual mismatch between detector pairs' beams may have to be more carefully diagnosed. In this work, we have measured and successfully contended with the dominant source of beam mismatch (differential pointing), but there may be other sources of mismatch in either the main beam, near sidelobes, or far sidelobes that could produce unknown levels of spurious polarization. New analysis techniques, aimed at directly convolving the CMB temperature with measured beam residuals, are currently under development.

Additional analysis is required to verify, and possibly change, the applied cuts and thresholds to verify the nominal BICEP2 data selection. By repeating a similar analysis to the work reported in Section 5.3.3, future analyses can demonstrate an insensitivity of measured bandpowers to small changes in the data selection, thereby verifying that highly aberrant data is either cut or appropriately down-weighted. Further data selection requirements may be unveiled by adding to BICEP1's jackknife arsenal. Some additional new data jackknives have already been introduced for BICEP2, but this too remains an area of active development for BICEP2.

Another requirement for a BICEP2 publication is a massive simulation run, containing $\sim$ 500 realizations of simulated signal and noise, derived from the real BICEP2 data. This simulation set is necessary for measuring the likelihood of various jackknife failures, estimating the noise and E-to-B

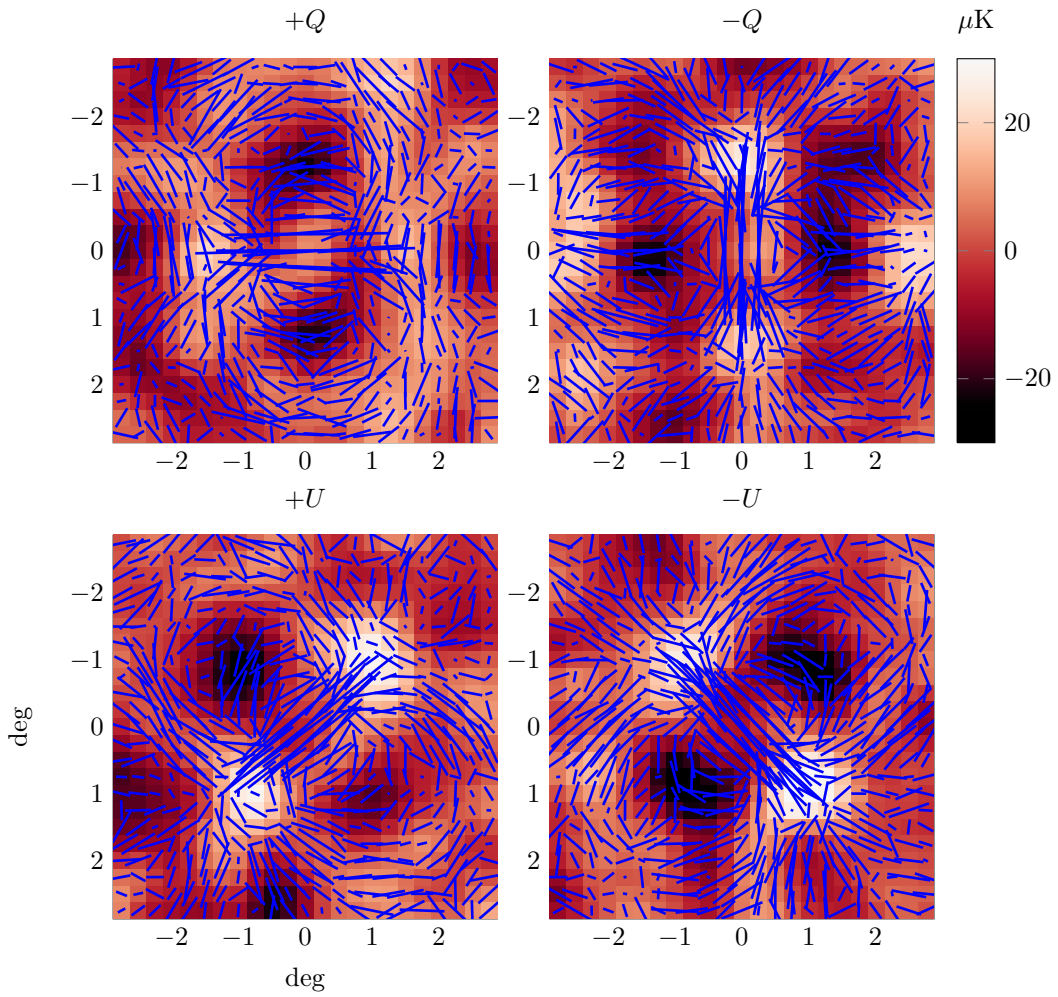


Figure 5.11: Stacked temperature anisotropies co-located at local maxima of $+Q$, $-Q$, $+U$, and $-U$ from three-year BICEP2 maps. Polarization vectors are overlaid on top of the temperature anisotropy plotted in the background image. We find that linear polarization in the CMB is directly tied to the amplitude and orientation of quadrupolar anisotropies. The quadrupole in $\pm Q$ is mildly distorted from filtering along the scan direction, which is aligned with the horizontal axis.

bias, measuring the filter function and suppression factor, and constructing likelihoods for parameter constraints.

While not strictly necessary, there are other analysis developments that are highly desirable for future BICEP2 analyses. Future BICEP2 analyses will greatly benefit from more sophisticated pseudo-$\mathcal{C}_\ell$ estimators. Owing to BICEP2's improved sensitivity, the sample variance from mixed E-modes into B contributed non-negligibly to BICEP2's total error budget. More sophisticated estimators that more efficiently separate E and B will result in lower sample variance, as compared to the BICEP1-style estimator.

With the advent of Planck, new and compelling joint analyses using both Planck temperature

and polarization data will be able to be pursued in the near future. Such analyses could be used to constrain foreground contamination in the BICEP2 maps. As we have previously remarked, since BICEP2 is a single band instrument, foreground component separation will have to rely on external data sets. Given Planck's sky coverage, frequency coverage, and angular resolution, it is ideally suited for this purpose.

Finally, future analyses may leverage the Keck Array as a cross-check against systematics. The Keck Array, fully deployed in 2011, is currently acquiring observations of the Southern Hole, BICEP2's chosen observing region. A future joint analysis with Keck offers not only added sensitivity, but also a powerful cross-check for instrument-induced systematics.

While much work remains, the fruits are enticing. BICEP2's sensitivity and thorough characterization, together with analysis methods successfully applied to BICEP1, promise in the very near future the best-ever measure of the B-mode polarization of the CMB, and a new observational window into the inflationary universe.

Appendix A

Calculations and digressions

A.1 Pair sum and difference signals

In this appendix, we step through the calculation of the per-detector pair-sum and pair-difference time series in the BICEP1 and BICEP2 analyses. Given a temperature anisotropy on the sky, we find the signal in the time series, given the instrument's beam, polarization, and frequency response, as well as a brightness fluctuation function. We let $\Delta T(\vec{p})$ represent the temperature anisotropy on the sky (for either the polarized or unpolarized field) as a function of position $\vec{p}$, the coordinate on either the celestial or equatorial sphere. The observed power fluctuation $\Delta P(\vec{p})$ seen by some detector is calculated as:

$$\Delta P(\vec{p}) = \int d\vec{p}' \int d\nu \Delta T(\vec{p}') \left(\frac{\tilde{B}(\vec{p} - \vec{p}', \nu)}{\tilde{\Omega}(\nu)} \right) \left(\frac{dI_\nu}{dT} A\tilde{\Omega}(\nu) F(\nu) \right). \tag{A.1}$$

Here, $\tilde{B}(\vec{p}, \nu)$ is the frequency-dependent beam function, $\tilde{\Omega}(\nu)$ is the frequency dependent beam solid angle, dI_ν/dT is the brightness fluctuation function, A is the aperture area, and $F(\nu)$ is the frequency bandpass. We define the beam function to be unitless, and thus it integrates to units of solid angle (requiring the normalization of $1/\tilde{\Omega}(\nu)$). The brightness fluctuation function dI_ν/dT has units of W/m^2/str/Hz/K, so the term $A\tilde{\Omega}(\nu)F(\nu)$ is necessary to convert into units of W/K. We can convert the power fluctuation back to an equivalent observed temperature fluctuation, given a position-independent brightness fluctuation to temperature conversion, evaluated at the central frequency of the band and approximated as constant across the bandpass. This conversion factor, $dI/dT_{\rm CMB}$, is calculated in Appendix A.2. The observed temperature fluctuation is thus:

$$\Delta T_{\rm obs}(\vec{p}) = \frac{\Delta P(\vec{p})}{dI/dT_{\rm CMB}}. \tag{A.2}$$

In practice, we cannot actually measure the frequency-dependent beam function $\tilde{B}(\vec{p}, \nu)$. Instead, what we measure is the response to some calibration source across the frequency bandpass of the

detector. In the limit that the calibration source has the same spectrum as the CMB, then the observed beam function, which we will call $B(\vec{p})$, is:

$$B(\vec{p}) = \frac{1}{dI/dT_{\rm CMB}} \int d\nu \left(\frac{\tilde{B}(\vec{p}-\vec{p}',\nu)}{\tilde{\Omega}(\nu)} \right) \left(\frac{dI_\nu}{dT} A\tilde{\Omega}(\nu) F(\nu) \right). \tag{A.3}$$

The prefactor of $1/(dI/dT_{\rm CMB})$ is required simply as a conversion factor to preserve units. Note that we have defined $B(\vec{p})$ to have units of inverse solid angle. The beam properties measured in far-field beam mapping will not be exactly equivalent to $B(\vec{p})$ because the thermal spectrum of the source is not matched to the CMB. However, this is a small correction. (In the case of BICEP2, the beam properties are measured with a chopped source, reflecting the ∼12 K zenith temperature at the South Pole.)

With the beam function defined in this way, we can write the response of detector i as a function of time. The i-th detector's pointing on the sphere as a function of time is denoted as $\vec{p}_i(t)$. The time series is then:

$$d_i(t) = g_i(t) \int d\vec{p}' B(\vec{p}_i(t) - \vec{p}') \Delta T(\vec{p}') + n_i(t). \tag{A.4}$$

Here, $g_i(t)$ is the responsivity to the CMB at 0 Hz (which we casually refer to as "gain") in units of ADU/μK,[1] and $n_i(t)$ is the noise. The true per-detector responsivity g_i is to be distinguished from the relative responsivity, which is defined below.

Now, rather than considering a generic temperature fluctuation $\Delta T(\vec{p}')$, we consider the temperature and polarization anisotropy fields, described by $\Theta(\vec{p}), Q(\vec{p})$, and $U(\vec{p})$:

$$d_i(t) = g_i(t) \int d\vec{p}' B(\vec{p}_i(t) - \vec{p}') \left[\Theta(\vec{p}') + \gamma_i (Q(\vec{p}') \cos 2\psi_i(t) + U(\vec{p}') \sin 2\psi_i(t)) \right] + n_i(t). \tag{A.5}$$

Here,ψ_i is the polarization orientation angle as illustrated in Figure 3.19, and γ_i is the polarization efficiency (as defined in Equation 4.4).

Next, we consider two detectors A and B within a polarized pair. We calculate the pair-sum and pair-difference TODs as:

$$d_s(t) = \frac{1}{2} \left(\frac{d_A(t)}{\eta_A} + \frac{d_B(t)}{\eta_B} \right) \tag{A.6}$$

$$d_f(t) = \frac{1}{2} \left(\frac{d_A(t)}{\eta_A} - \frac{d_B(t)}{\eta_B} \right). \tag{A.7}$$

Here, η_A and η_B are the applied relative responsivity corrections. In the case of BICEP1 and BICEP2, these are derived from elevation nods (or "el-nods", as in Section 4.1.1). The relative responsivity

[1] ADU, or analog-to-digital units, is the unit used by the data acquisition system to report the voltage of the detector. ADU can be converted to Volts with a single calibration factor.

correction is calculated as:

$$\eta_i = \frac{g_{e,i}}{\langle g_e \rangle}. \tag{A.8}$$

Here, $g_{e,i}$ is the responsivity of the i-th detector to an el-nod, and $\langle g_e \rangle$ is the median el-nod-derived responsivity over all detectors on the focal plane. Both have units of ADU per airmass, and, as a result, η_i is unitless. The el-nod corrected responsivity is thus:

$$\tilde{g}_i = \frac{g_i}{\eta_i} \left[\frac{\mathrm{ADU}}{\mu\mathrm{K}}\right]. \tag{A.9}$$

Let us now consider the idealized case, wherein detectors within a polarized pair have perfectly matched beams, pointing centroids, and relative gains. We let $B_A(\vec{p}) = B_B(\vec{p}) = B_0(\vec{p})$, $\vec{p}_A(t) = \vec{p}_B(t) = \vec{p}_0(t)$, and $\tilde{g}_A = \tilde{g}_B = \tilde{g}$. We also assume perfect polarization orthogonality and efficiency, setting $\psi_A = \psi$, $\psi_B = \psi + \pi/2$, and $\gamma = 1$. The pair-sum and difference is then constructed as:

$$d_s(t) = \tilde{g} \int d\vec{p}\,' B_0(\vec{p}_0(t) - \vec{p}\,')\Theta(\vec{p}\,') = \tilde{g}(B_0 * \Theta)(\vec{p}_0(t)) + n_s(t) \tag{A.10}$$

$$d_f(t) = \tilde{g}[B_0 * (Q\cos 2\psi + U \sin 2\psi)](\vec{p}_0(t)) + n_d(t). \tag{A.11}$$

Here, $n_s(t)$ and $n_d(t)$ are the pair-sum and difference noise components, respectively. In Equation A.10, we have written the integral as a convolution of the temperature field with the beam. For notational simplicity, we define $d_T(t)$ and $d_P(t)$ to be the noiseless pair-sum and difference signal from the true sky temperature and polarization, respectively:

$$d_T(t) = \tilde{g}(B_0 * \Theta)(\vec{p}_0(t)) \tag{A.12}$$

$$d_p(t) = \tilde{g}[B_0 * (Q\cos 2\psi + U \sin 2\psi)](\vec{p}_0(t)). \tag{A.13}$$

A.2 CMB surface brightness

In this appendix we calculate the relationship between temperature fluctuations on the sky to optical power fluctuations seen by a detector. In the Rayleigh-Jeans limit, the intensity of a blackbody spectrum with temperature T_{RJ} is:

$$I_\nu = 2kT_{\mathrm{RJ}}\frac{\nu^2}{c^2} \left[\frac{\mathrm{W}}{\mathrm{m}^2\ \mathrm{Hz\ sr}}\right], \tag{A.14}$$

or, integrated over a finite bandwidth (to first order in $\Delta\nu$):

$$I \simeq 2kT_{\mathrm{RJ}}\nu^2\frac{\Delta\nu}{c^2} \left[\frac{\mathrm{W}}{\mathrm{m}^2\ \mathrm{sr}}\right]. \tag{A.15}$$

To calculate total intensity (instead of flux), we need to calculate $A\Omega_{\rm pix}$, the solid angle subtended by a given detector multiplied by the effective light-collecting area. Since our detectors are single-moded, this is just equal to λ^2, up to factors of order unity. Here's why: single moded optical systems preserve $A\Omega$, which means that rather than calculating this quantity at the aperture (which is hard, because of the non-trivial illumination of the primary optic), we can calculate the quantity at the antenna. This is much easier, as the illumination pattern of the antenna is a top-hat. Differences can be accounted for by an overall efficiency factor η, which includes the total throughput.

$\Omega_{\rm pix}$ can be easily calculated for a top-hat of width w at a wavelength λ with the well-known formula:

$$\Omega = \left(1.44\frac{\lambda}{w}\right)^2. \tag{A.16}$$

$A\Omega$ is then simply $(1.44\lambda)^2$.

Returning to our brightness fluctuation, we multiply by $A\Omega$ for a single detector and an overall optical efficiency factor $\eta = 0.4$ and take the trivial derivative with respect to $T_{\rm RJ}$:

$$\frac{dI}{dT_{\rm RJ}} \simeq 2k\nu^2\eta A\Omega\frac{\Delta\nu}{c^2} = 4.14k\eta\Delta\nu \simeq 8\times10^{-13}\ \left[\frac{W}{K_{\rm RJ}}\right]. \tag{A.17}$$

To convert to $T_{\rm CMB}$, we multiply by a factor of 1.7, which is the ratio of the integral of a 3 K blackbody over this band to the Rayleigh-Jeans approximation. We finally arrive at the brightness temperature fluctuation relation:

$$\frac{dI}{dT_{\rm CMB}} \simeq 1\times10^{-12}\ \left[\frac{W}{K_{\rm CMB}}\right]. \tag{A.18}$$

Appendix B

Dielectric sheet calibrator model

Deriving precise polarization angle measurements using the dielectric sheet calibrator described in Section 3.4.1 requires an accurate model of the reflected and transmitted power through the beam splitter as a function of several model parameters. We describe that model below, following closely the model presented in (Takahashi 2010).

We will define the following local coordinate system in which to establish our model: $+z$ increases along the boresight toward the sky with $z = 0$ corresponding to the plane of the detectors. The positive y axis is aligned with $\theta = 0$, as in Figure 3.19, with DK increasing CCW. This puts detector tile 1 in the $+x, +y$ quadrant of our coordinate system. We work in a coordinate frame in which the focal plane is fixed, so that polarization modulation results from motion of the dielectric sheet relative to the coordinate frame. We let ϕ represent the rotational position of the dielectric sheet with respect to the focal plane, equal to the DK angle of the telescope, with zero defined as usual. As a practical matter, we use the telescope command coordinates because we are interested in the DK angle with respect to the ground, rather than the telescope topocentric horizontal coordinates (as in Equation C.3). Usually, these are very close to identical, but since this measurement is taken very near the coordinate singularity at zenith, the discrepancy between them can be large.

The tilt t of the dielectric sheet is measured from horizontal. This is nominally 39 degrees plus the measured tilt at the start of the scan, which is typically close to 6.4 degrees. The dielectric sheet is made from Mylar, which has a nominal index $n = 1.83$ (Lamb 1996). The BICEP2 data were taken with sheet thicknesses of 1 and 2 mil, measured in the lab with a micrometer. We assume a nominal band center $\lambda = 2.07$ mm (145 GHz).

Calculating the normal vector of the dielectric sheet requires some of the same geometric machinery used in Appendix C. We begin with the calibrator sheet's normal vector co-aligned with $\hat{\mathbf{z}}$,

and then apply the following rotation matrices:

$$\mathrm{R_t} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{bmatrix}, \mathrm{R}_\gamma = \begin{bmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{bmatrix}, \mathrm{R}_\phi = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

R_ϕ is the rotation of the calibrator with respect to the focal plane due to DK rotation, R_γ accounts for a potential lateral tilt[1] of the dielectric sheet (typically small), and $\mathrm{R_t}$ accounts for the tilt angle of the dielectric sheet, typically close to 45 degrees. These rotation matrices obviously do not commute. The correct sequence of rotations can be applied[2] to calculate the normal vector:

$$\hat{n} = \hat{n}_0 \mathrm{R}_t \mathrm{R}_\gamma \mathrm{R}_\phi. \tag{B.1}$$

In vector notation, this reduces to:

$$\hat{n} = (\cos\phi\sin\gamma\cos t - \sin t\sin\phi)\hat{x} + (-\cos t\sin\phi\sin\gamma + \sin t\cos\phi)\hat{y} + \cos t\cos\gamma\hat{z}. \tag{B.2}$$

The angle of incidence of each detector's beam can be calculated by considering the ray of each detector as it exits the aperture. This is parameterized by r_p and θ_p, the radial distance of the detector's centroid from the boresight and its bearing, respectively. The bearing angle θ_p increases CCW from the x axis of our coordinate system. These are equivalent to the beam centroids as measured from the boresight. The resulting ray is:

$$\hat{b}(r_p, \theta_p) = \sin(r_p)\cos(\theta_p)\hat{x} + \sin(r_p)\sin(\theta_p)\hat{y} + \cos(r_p)\hat{z}. \tag{B.3}$$

We calculate the angle of incidence with the dielectric sheet as:

$$\theta_i = \arccos(\hat{b}\cdot\hat{n}). \tag{B.4}$$

To simplify the notation, we introduce θ_n, the angle of incidence within the dielectric:

$$\theta_n = \arcsin(\sin(\theta_i)/n). \tag{B.5}$$

Similarly, we define:

$$\beta = \frac{4\pi n d}{\lambda}\cos(\theta_n). \tag{B.6}$$

[1] In the BICEP2 observation logbook, the sign of the measured lateral tilt is recorded with a forward or backslash. In our definition of γ, a backslash corresponds to positive γ, while a forward slash corresponds to negative γ.

[2] One might be tempted to switch the rotations of t and γ, but doing so would be incorrect: if we slew the telescope to a new elevation, we will still measure the same lateral tilt angle γ.

The last piece we need is the thing we are trying to measure: the axes of sensitivity of detectors A and B. We parameterize the polarization angle, by which we mean the axes of maximum sensitivity, by χ. In the case that $\chi = 0$, the polarization axes will be co-aligned with the vector that connects the boresight and θ_p. The polarization angle χ is therefore referenced from θ_p. We represent the orientation of the linear polarization axes for A and B with a "headed" vector, expressed in our local coordinate frame as:

$$\hat{A}(\chi) = (-\sin(\chi)\cos(r_p)\cos(\theta_p) - \cos(\chi)\sin(\theta_p))\hat{x} \tag{B.7}$$

$$+ \cos(\theta_p)\cos(\chi) - \sin(\chi)\cos(r)\sin(\theta)\hat{y} \tag{B.8}$$

$$+ \sin(\chi)\sin(r)\hat{z} \tag{B.9}$$

$$\hat{B}(\chi) = (\cos(\chi)\cos(r_p)\cos(\theta_p) - \sin(\chi)\sin(\theta_p))\hat{x} \tag{B.10}$$

$$+ \cos(\theta_p)\sin(\chi) + \cos(\chi)\cos(r)\sin(\theta)\hat{y} \tag{B.11}$$

$$- \cos(\chi)\sin(r)\hat{z}. \tag{B.12}$$

Note that in this step we have assumed perfect orthogonality between A and B, so that the polarization axes can be described by a single angle, rather than fitting the two angles separately. We make this assumption because of the large common-mode atmospheric contamination in A and B. By performing a fit to the pair-difference signal rather than individually fitting A and B, we can reject the variable atmospheric signal and vastly improve the fit. We also note that while we have represented the polarization pair with the single angle χ, we can readily recover the polarization angles of A and B:

$$\chi_A = \chi \tag{B.13}$$

$$\chi_B = \chi + \pi/2. \tag{B.14}$$

For the sake of clarity, we will also introduce an angle α to visualize deviations of the polarization angle from nominal. We define this angle to be:

$$\alpha = -(\chi + \theta_p) + 90. \tag{B.15}$$

In this definition, α will be close to 0 for all A detectors, and close to ± 90 for all B detectors.

We can now calculate the coupling fraction of polarizations A and B to the reflected TE mode:

$$f_{\mathrm{A,TE}} = \hat{A} \cdot \frac{\hat{b} \times \hat{n}}{|\hat{b} \times \hat{n}|} \tag{B.16}$$

$$f_{\mathrm{B,TE}} = \hat{B} \cdot \frac{\hat{b} \times \hat{n}}{|\hat{b} \times \hat{n}|}. \tag{B.17}$$

From orthogonality, the coupling fraction to the TM mode can be calculated from the coupling to the TE mode as:

$$f_{\mathrm{A,TM}} = \sqrt{1 - f_{\mathrm{A,TE}}^2} \tag{B.18}$$

$$f_{\mathrm{B,TM}} = \sqrt{1 - f_{\mathrm{B,TE}}^2}. \tag{B.19}$$

As computed in Takahashi 2010, we can find the reflection coefficient of the TE and TM modes by a superposition of the reflections that occur upon entering and exiting the dielectric material:

$$R_{\mathrm{TE,TM}} = \frac{r_1^2 + r_2^2 + 2r_1 r_2 \cos\beta}{1 + r_1^2 r_2^2 + 2r_1 r_2 \cos\beta}. \tag{B.20}$$

Here, r_1 and r_2 are the reflection coefficients from entering and exiting the dielectric material, calculated as:

$$r_{1,\mathrm{TE}} = \frac{\cos\theta_i - n\cos\theta_n}{\cos\theta_i + n\cos\theta_n}, \quad r_{2,\mathrm{TE}} = -r_{1,\mathrm{TE}} \tag{B.21}$$

$$r_{1,\mathrm{TM}} = \frac{n\cos\theta_i - \cos\theta_n}{n\cos\theta_i + \cos\theta_n}, \quad r_{2,\mathrm{TM}} = -r_{1,\mathrm{TM}}. \tag{B.22}$$

The reflected modes seen by detectors A and B can finally be calculated:

$$R_A = R_{\mathrm{TE}} f_{\mathrm{A,TE}}^2 + R_{\mathrm{TM}} f_{\mathrm{A,TM}}^2 = R_{\mathrm{TM}} + (R_{\mathrm{TE}} - R_{\mathrm{TM}}) f_{\mathrm{A,TE}}^2 \tag{B.23}$$

$$R_B = R_{\mathrm{TE}} f_{\mathrm{B,TE}}^2 + R_{\mathrm{TM}} f_{\mathrm{B,TM}}^2 = R_{\mathrm{TM}} + (R_{\mathrm{TE}} - R_{\mathrm{TM}}) f_{\mathrm{B,TE}}^2. \tag{B.24}$$

By pair differencing, we reject the transmitted modes through the beam splitter, and we are left with the purely polarized intensity, which is:

$$I_{\mathrm{diff}} = \Delta I (R_A - R_B). \tag{B.25}$$

This model thus reduces to two free parameters: the angle of the polarization axes, α, and an amplitude that corresponds to the effective brightness difference between the warm absorber surrounding the dielectric sheet and the sky, ΔI.

Appendix C

Three-dimensional pointing model for BICEP2 beam mapping

In this appendix, we document a fully general pointing model used for analyzing beam maps taken at the South Pole. BICEP2 employs a folding flat mirror to redirect the main beam to a point near the horizon, which sits 2.98 m above the intersection of the AZ, EL, and DK rotation axes. With a source distance of ~ 200 m, the movement of the mirror introduces parallax effects that must be taken into account.

Our approach differs somewhat from the offline pointing model, used to reconstruct the boresight position on the sky using fits derived from star pointing. There, the analysis focuses on a variety of coordinate transformations, given various tilt, offset, and flexure terms. Rather than approaching the problem as a series of coordinate transformations, we instead take a "ray tracing" approach, treating the boresight as a vector in a fixed coordinate system. We represent each of the detectors' chief rays with a vector, and then perform a series of operations on this "ray bundle" to recover the projected elevation and azimuth of the detector centroid on a sphere with a radius equal to the source distance. The origin of our coordinate system is the intersection of the elevation, azimuthal, and boresight axes, which we assume to be coincident. This is a good approximation for BICEP2, but not necessarily so for Keck. However, this can be accounted for by simple vector addition.

We use the same definitions as in Yoon 2007 for a few distinct coordinate systems that will be referenced:

$$[AZ, EL, DK]_0, \text{ Raw encoder counts} \tag{C.1}$$

$$[AZ, EL, DK]_c, \text{ Command coordinates} \tag{C.2}$$

$$[AZ, EL, DK], \text{ Topocentric ideal horizontal coordinates.} \tag{C.3}$$

C.1 Raw encoder counts, command coordinates, and boresight pointing

We begin by noting the conversion from raw encoder counts to command coordinates with and without the mirror for BICEP2:

$$[AZ, EL, DK]_c = ([AZ, EL, DK]_0 + \texttt{encoder\_zero[0,1,2]})(\texttt{encoder\_mul[0,1,2]}). \quad \text{(C.4)}$$

In the presence of the mirror, the command coordinates must be redefined. Specifically, the encoder zeroes and the encoder multipliers, which define the sense of positive and negative motion, are as follows with and without the mirror (as defined in telescope's schedule library and beam mapping schedule, respectively):

$$\texttt{encoder\_zeros} = [-322.49, -82.44, 74] \quad \text{(C.5)}$$

$$\texttt{encoder\_zeros}\,(\text{with mirror}) = [-143.4, 1.2, -74] \quad \text{(C.6)}$$

$$\texttt{encoder\_mul} = [2304000, 2304000, 574400] \quad \text{(C.7)}$$

$$\texttt{encoder\_mul}\,(\text{with mirror}) = [2304000, -2304000, -574400]. \quad \text{(C.8)}$$

We compute the topocentric, ideal horizontal coordinates of the boresight $\vec{b}$ using the usual offline pointing model for BICEP2, the details of which can be found in Yoon 2007, Chiang 2008, and references therein. In the ray tracing pointing model, we represent the boresight as the vector $\vec{b}$, extending from the origin of the coordinate system to the intersection with the folding flat mirror (Figure C.1). The magnitude is equal to the mirror height, $|\vec{b}| = 2.98$ m.

C.1.1 Calculating the mirror normal vector

Next, we calculate $\hat{n}$, the mirror normal vector. The mirror position is constrained by considering 5 parameters (some of which are degenerate): the telescope command azimuth and elevation, $[AZ, EL]_c$ (as the mirror moves with the telescope), the mirror tilt, the mirror roll, and the mirror height. We ignore azimuth and elevation tilt terms here and instead account for them with a small adjustment to the fitted mirror tilt and roll. The mirror tilt is degenerate with EL_c, while mirror roll is degenerate with AZ_c. We treat them separately, as the mirror tilt and roll are assumed to be constant in time, while the $[AZ, EL]_c$ are changing as the telescope scans. The mirror tilt and roll are both measured with respect to gravity from horizontal. In November 2011 we measured the mirror tilt to be 40.9 degrees and the roll to be close to zero. We calculate the mirror normal vector by performing a series of Euler rotations on a unit vector, illustrated in Figure C.2. The Euler angles are calculated as in Equations C.9 through C.12. (For BICEP2, $\alpha = 180 - AZ$ because the mirror is

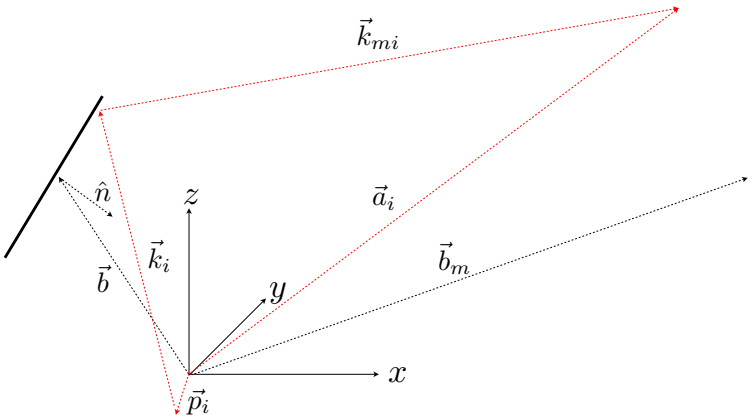


Figure C.1: Cartoon of the full pointing model. The goal is to find $\vec{a}_i$, the apparent detector centroid position of the ith detector pair. The vectors $\vec{b}$, $\vec{b}_m$, and $\hat{n}$ represent the boresight, the reflected boresight, and the mirror normal vector, respectively. The vector $\vec{p}_i$ is the physical ith pixel location on the focal plane at this telescope position, and $\vec{k}_i$ the detector's chief ray. The vector $\vec{k}_{mi}$ is the result of reflecting $\vec{k}$ about the plane defined by $\hat{n}$.

installed to reflect the beam back in the opposite direction as AZ and is independent of DK).

$$\alpha = 180 - AZ_c \tag{C.9}$$

$$\beta = \pi + (\text{tilt}) - EL_c \tag{C.10}$$

$$\gamma = (\text{roll}) \tag{C.11}$$

$$\hat{n} = \hat{n}_0(\mathrm{R}_\gamma \mathrm{R}_\beta \mathrm{R}_\alpha). \tag{C.12}$$

The subsequent Euler rotation matrices are thus:

$$\mathrm{R}_\alpha = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathrm{R}_\beta = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \mathrm{R}_\gamma = \begin{bmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

To calculate the reflected boresight $\hat{b}_m$, we insist that the component of the boresight unit vector parallel to the mirror be zero:

$$\vec{b}_m = r(\hat{b} - 2(\hat{b} \cdot \hat{n})\hat{n}). \tag{C.13}$$

C.1.2 Determining the physical detector location

To produce a faithful ray-trace representation of the optical response of each detector, we must determine the physical detector location, which will serve as the origin of each ray. We first calculate

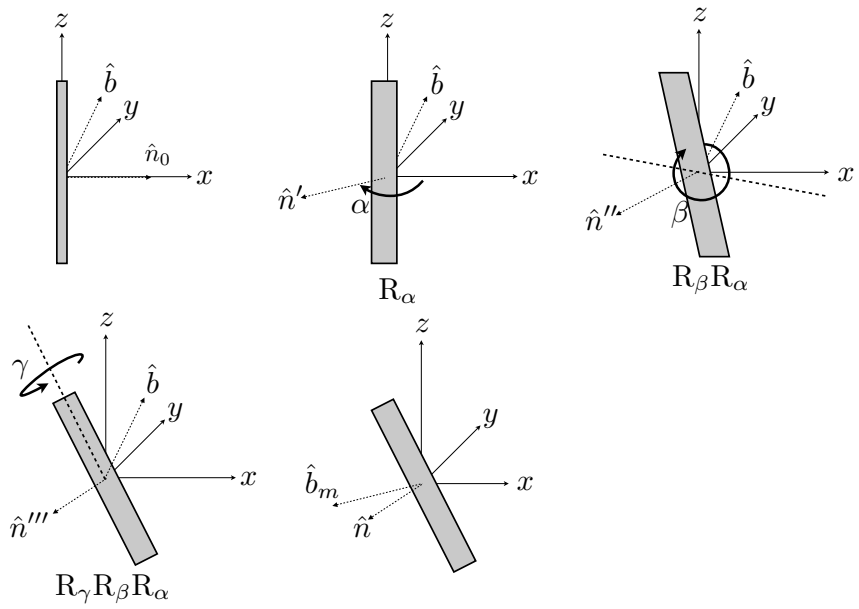


Figure C.2: Graphical representation of the mirror normal vector rotation and boresight reflection. We rotate the normal vector by the Euler angles α, β, γ. The values of β and γ can be either measured with a protractor or constrained using multiple DK angle beam maps.

the physical position of each detector on the focal plane when the telescope is at $[AZ, EL, DK] = [0, 90, 0]$. For BICEP2, this is just calculated using the plate scale (in degrees/mm) and a Cartesian flat plane conversion of $[r, \theta]$ to x and y (i.e. $r\cos(\theta)$, $r\sin(\theta)$). We assume that the center of the focal plane is located at the origin of the coordinate system, which, for BICEP2 is an excellent approximation. We denote the vector that connects the origin to the detector's physical location at [0, 90, 0] as $\vec{p}'_{0i}$. To calculate the position at some arbitrary commanded position, we rotate $\vec{p}'_{0i}$ by $[AZ, EL, DK]_c$ (the current commanded position), to find $\vec{p}_i$, the physical location of the pixel at this telescope position. The reason to use the command position rather than the offline pointing model is because we wish to calculate the physical location relative to the origin, rather than a position relative to the celestial sphere.

C.1.3 Calculating the chief ray and reflected chief ray

Next, we calculate $\hat{k}_i$, the unit vector representing the beam centroid of detector pair i. (This is equivalent to a chief ray in traditional ray tracing). We calculate this as a reckoning from the boresight. Notice that, unlike the physical detector location, we calculate $\hat{k}$ with respect to the optical boresight, $[AZ, EL]$ (as derived from star-pointing), rather than the command coordinates.

Next, we find the point at which this ray bundle $\hat{k}$ intersects the plane of the mirror defined by

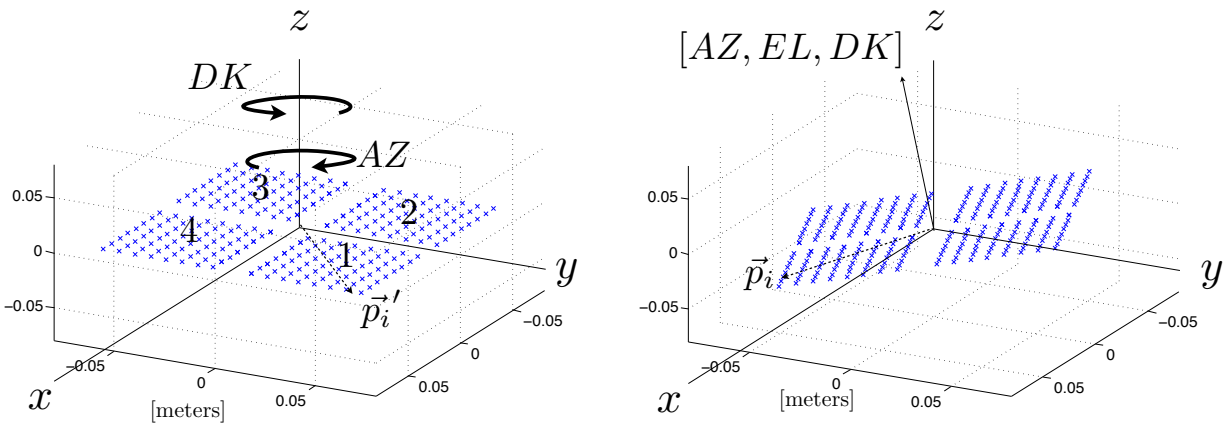


Figure C.3: Physical location of all BICEP2 detectors in meters when the telescope is at $[0, 90, 0]$ (left) and at the boresight position $[AZ, EL, DK]$ (right). The sense of increasing AZ and DK is shown at left.

$\hat{n}$. We will denote this as $\vec{k}$. (Here $\vec{p}$ represents the ray bundle of all $\vec{p}_i$).

$$\vec{k} = \frac{(\vec{c} - \vec{p}) \cdot \hat{n}}{\hat{k} \cdot \hat{n}} \hat{k}. \tag{C.14}$$

C.1.4 Calculating the apparent detector centroid

Finally, we can calculate $\vec{a}_i$, the apparent boresight position as seen by the ith detector pair. This can be calculated from simple vector addition, as we can see from Figure C.1:

$$\vec{a}_i = \vec{p}_i + \vec{k}_i + \vec{k}_{mi}. \tag{C.15}$$

The result is shown for two telescope positions in Figure C.5.

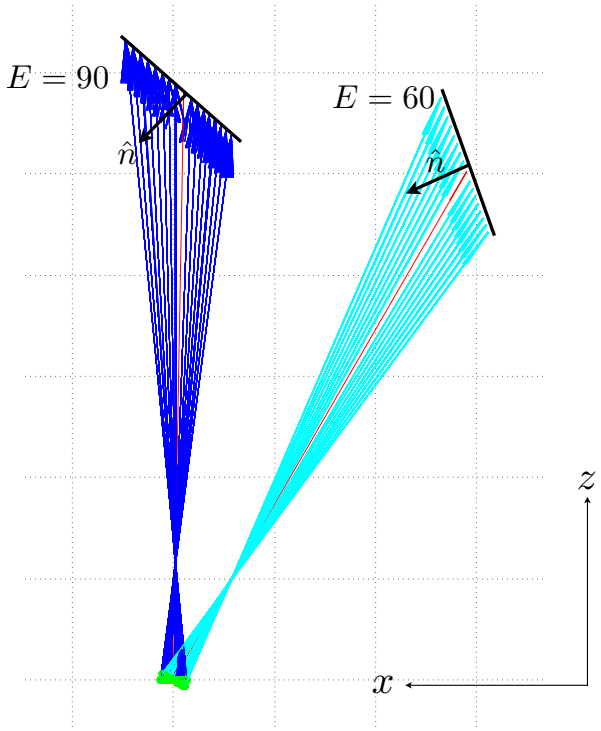


Figure C.4: Ray bundle $\vec{k}$ for two different telescope elevations, $EL = 90$ and $EL = 60$. the red vectors represent the nominal boresight position $\vec{b}$, and the black vectors represent $\hat{n}$, the normal vector to the mirror face. The green points represent $\vec{p}_i$, the physical pixel locations.

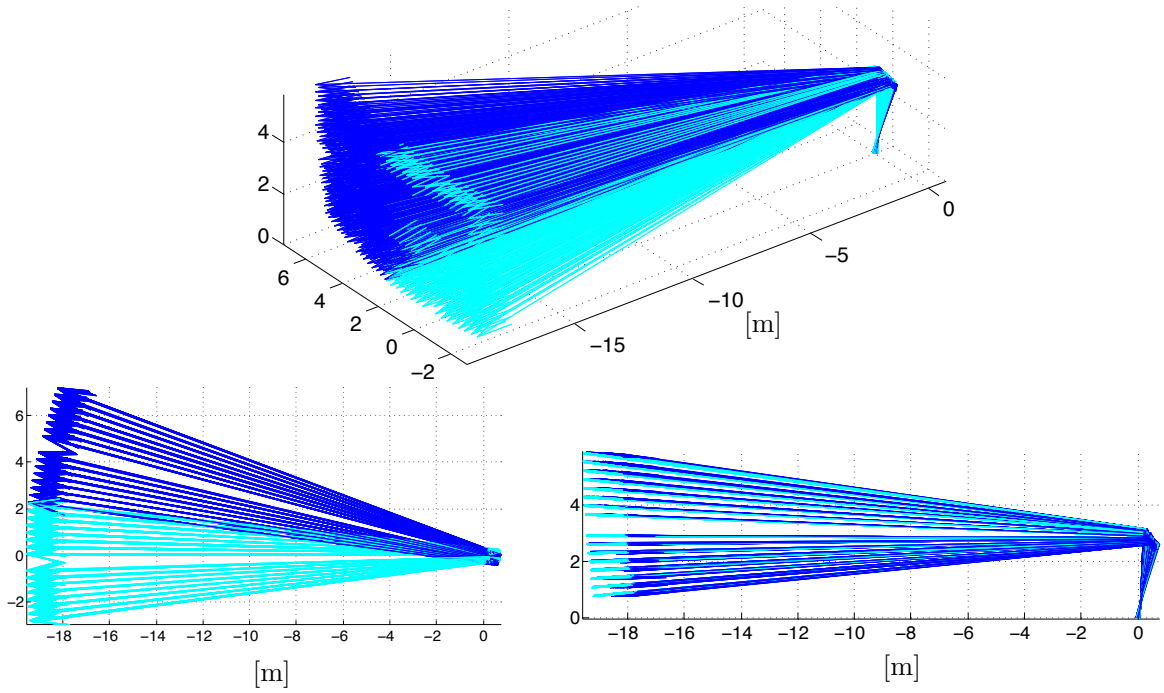


Figure C.5: Full ray tracing model for $[AZ, EL, DK] = [0, 80, 0]$ (cyan) and $[AZ, EL, DK] = [15, 80, 0]$ (blue). Each line traces the chief ray of each detector from the focal plane to the final source distance (here we have calculated the ray tracing assuming a source distance of 20 meters to exaggerate the effect).

Appendix D

Measured properties and operational notes for BICEP2 detectors

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
0	A	1	RGL	5.74	81.27	9.61	0.229	0.065	0.003
1	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
2	B	2	RGL	6.65	75.82	105.18	0.228	-0.034	-0.001
3	A	2	RGL	6.66	75.70	15.18	0.225	-0.012	0.003
4	B	3	RGL	7.21	63.68	117.34	0.226	-0.051	0.031
5	A	3	RGL	7.22	63.57	27.34	0.222	-0.034	0.024
6	B	4	RGL	8.04	53.67	127.33	0.228	-0.056	0.074
7	A	4	RGL	8.05	53.65	37.33	0.228	-0.048	0.056
8	B	5	Dark	9.62	46.88	135.23	0.247	-0.015	-0.080
9	A	5	Dark	8.90	44.32	45.23	0.295	0.138	0.151
10	B	6	RGL	8.56	49.49	131.52	0.235	-0.037	0.117
11	A	6	RGL	8.56	49.36	41.52	0.230	-0.021	0.083
12	B	7	RGL	7.59	58.52	122.61	0.228	-0.065	0.048
13	A	7	RGL	7.59	58.33	32.61	0.226	-0.049	0.045
14	B	8	RGL	6.89	69.54	111.51	0.224	-0.047	0.012
15	A	8	RGL	6.90	69.45	21.51	0.222	-0.030	0.012
16	B	9	RGL	6.51	82.51	98.45	0.238	0.042	-0.022
17	A	9	RGL	6.52	82.32	8.45	0.238	0.033	0.009
18	B	10	RGL	5.88	74.07	107.18	0.236	0.091	-0.003
19	A	10	RGL	5.90	73.93	17.18	0.237	0.104	0.007
20	B	11	RGL	6.50	60.78	120.50	0.219	-0.009	0.017
21	A	11	RGL	6.51	60.68	30.50	0.218	-0.007	0.004
22	B	12	RGL	7.41	50.19	130.99	0.222	0.003	0.052
23	A	12	RGL	7.40	50.14	40.99	0.222	0.005	0.041
24	B	13	RGL	8.53	42.06	138.96	0.230	0.021	0.113
25	A	13	RGL	8.54	41.98	48.96	0.231	0.033	0.099
26	B	14	RGL	7.96	45.79	135.25	0.225	-0.001	0.078
27	A	14	RGL	7.96	45.76	45.25	0.224	0.003	0.053
28	B	15	Flagged	6.94	56.16	126.09	0.220	-0.004	0.029
29	A	15		6.96	55.35	36.09	0.220	-0.011	0.029
30	B	16	RGL	6.16	66.95	114.19	0.218	-0.009	0.004
31	A	16	RGL	6.18	66.83	24.19	0.218	-0.005	0.008
32	B	1	RGL	5.73	81.40	99.61	0.228	0.045	-0.011
33	A	17		4.23	77.68	13.22	0.221	-0.023	0.004
34	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
35	B	18	RGL	5.15	71.76	109.77	0.221	0.010	0.002

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
36	A	18	RGL	5.16	71.64	19.77	0.219	0.006	0.008
37	B	19	RGL	5.84	57.01	124.40	0.218	-0.003	0.006
38	A	19	RGL	5.85	56.90	34.40	0.218	-0.007	0.004
39	B	20	RGL	6.82	46.09	135.28	0.222	0.003	0.039
40	A	20	RGL	6.83	46.06	45.28	0.221	0.007	0.028
41	B	21	RGL	8.02	37.90	143.17	0.226	0.024	0.087
42	A	21	RGL	8.03	37.83	53.17	0.228	0.028	0.068
43	B	22	Flagged	7.46	40.46	139.54	0.220	0.000	0.000
44	A	22		7.46	40.46	49.54	0.221	0.018	0.039
45	B	23	RGL	6.32	51.08	130.28	0.220	0.000	0.019
46	A	23	RGL	6.32	51.05	40.28	0.220	0.004	0.006
47	B	24	RGL	5.46	63.75	117.57	0.220	-0.013	0.005
48	A	24	RGL	5.48	63.56	27.57	0.224	-0.029	-0.011
49	B	25	RGL	4.97	79.90	101.13	0.225	0.034	-0.002
50	A	25	RGL	4.99	79.75	11.13	0.228	0.065	0.006
51	B	26	RGL	4.43	68.46	113.22	0.219	0.007	-0.005
52	A	26	RGL	4.45	68.30	23.22	0.221	-0.005	0.007
53	B	27	RGL	5.22	52.44	129.26	0.217	0.009	0.008
54	A	27	RGL	5.23	52.38	39.26	0.216	0.000	0.004
55	B	28	RGL	6.30	41.19	140.32	0.219	0.013	0.023
56	A	28	RGL	6.31	41.12	50.32	0.218	0.006	0.017
57	B	29	RGL	7.58	33.30	147.89	0.222	0.032	0.052
58	A	29	RGL	7.59	33.24	57.89	0.225	0.043	0.044
59	B	30	RGL	6.89	36.95	144.46	0.221	0.014	0.032
60	A	30	RGL	6.90	36.91	54.46	0.221	0.012	0.024
61	B	31	RGL	5.73	46.15	135.32	0.217	0.004	0.013
62	A	31	RGL	5.75	46.12	45.32	0.217	0.002	0.005
63	B	32	RGL	4.79	59.63	121.93	0.223	-0.005	0.005
64	A	32	RGL	4.81	59.53	31.93	0.218	-0.003	0.000
65	B	17	Flagged	4.22	77.93	103.22	0.222	0.023	-0.000
66	A	33		2.74	70.44	20.97	0.224	0.041	0.017
67	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
68	B	34	RGL	3.74	63.58	118.02	0.220	0.011	-0.003
69	A	34	RGL	3.76	63.45	28.02	0.219	0.008	0.009
70	B	35	RGL	4.64	46.50	135.38	0.215	-0.005	0.007
71	A	35	RGL	4.67	46.45	45.38	0.217	-0.001	0.019
72	B	36	RGL	5.82	35.29	146.22	0.217	0.008	0.026
73	A	36	RGL	5.83	35.33	56.22	0.217	0.008	0.026
74	B	37	RGL	7.19	28.03	153.16	0.222	0.040	0.054
75	A	37	RGL	7.21	28.02	63.16	0.223	0.047	0.040
76	B	38	RGL	6.46	31.39	150.05	0.219	0.026	0.030
77	A	38	RGL	6.47	31.37	60.05	0.219	0.020	0.022
78	B	39	RGL	5.19	40.29	141.44	0.217	-0.000	0.007
79	A	39	RGL	5.21	40.21	51.44	0.217	-0.004	0.005
80	B	40	Flagged	2.95	109.71	127.70	0.220	0.000	0.000
81	A	40		4.21	53.16	37.70	0.231	-0.046	0.100
82	B	41	RGL	3.46	74.99	106.24	0.220	0.000	0.000
83	A	41	RGL	3.49	74.85	16.24	0.220	0.011	0.010
84	B	42	RGL	3.07	57.23	125.00	0.220	0.014	0.001
85	A	42	RGL	3.09	57.11	35.00	0.218	0.006	0.012
86	B	43	RGL	4.11	39.23	143.14	0.217	-0.002	0.008
87	A	43	RGL	4.13	39.15	53.14	0.216	-0.001	0.008
88	B	44	RGL	5.40	28.70	153.06	0.218	0.014	0.016
89	A	44	RGL	5.42	28.73	63.06	0.217	0.004	0.010

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
90	B	45	RGL	6.83	22.38	158.97	0.222	0.044	0.037
91	A	45	RGL	6.86	22.41	68.97	0.223	0.048	0.017
92	B	46	RGL	6.11	25.12	156.36	0.214	0.038	0.027
93	A	46	RGL	6.12	25.19	66.36	0.220	0.020	0.017
94	B	47	RGL	4.74	33.09	148.79	0.216	0.001	0.011
95	A	47	RGL	4.76	32.99	58.79	0.217	0.011	0.007
96	B	48	RGL	3.53	46.59	135.48	0.219	0.007	0.004
97	A	48	RGL	3.55	46.57	45.48	0.217	0.003	0.013
98	B	33	Flagged	2.72	70.61	110.97	0.225	0.040	-0.001
99	A	49		1.39	49.44	46.17	0.259	-0.005	0.002
100	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
101	B	50	RGL	2.45	47.43	135.67	0.232	0.086	-0.007
102	A	50	RGL	2.48	47.38	45.67	0.225	0.055	0.015
103	B	51	RGL	3.66	29.78	152.85	0.216	0.013	0.001
104	A	51	RGL	3.68	29.92	62.85	0.216	0.004	0.002
105	B	52	RGL	5.09	21.12	160.83	0.217	0.018	0.013
106	A	52	RGL	5.11	21.21	70.83	0.217	0.011	0.007
107	B	53	RGL	6.57	16.14	165.27	0.224	0.058	0.034
108	A	53	RGL	6.58	16.18	75.27	0.225	0.053	0.016
109	B	54	RGL	5.83	18.34	163.33	0.220	0.035	0.023
110	A	54	RGL	5.85	18.39	73.33	0.219	0.022	0.009
111	B	55	RGL	4.37	24.64	157.49	0.217	0.007	0.010
112	A	55	RGL	4.39	24.67	67.49	0.217	0.000	0.005
113	B	56	RGL	3.02	36.69	146.08	0.215	0.003	0.004
114	A	56	RGL	3.04	36.79	56.08	0.219	0.013	0.008
115	B	57	RGL	2.02	63.64	119.27	0.229	0.037	0.049
116	A	57	RGL	2.04	63.64	29.27	0.222	0.029	0.011
117	B	58		1.97	27.73	152.28	0.220	-0.000	0.002
118	A	58	Flagged	1.97	27.73	62.27	0.220	0.000	0.000
119	B	59	RGL	3.34	18.01	164.58	0.215	-0.021	0.003
120	A	59	RGL	3.37	18.21	74.58	0.217	-0.011	0.006
121	B	60	RGL	4.84	12.45	169.41	0.220	-0.008	0.011
122	A	60	RGL	4.86	12.57	79.41	0.219	-0.009	0.000
123	B	61	Dark	6.20	5.68	171.95	0.220	0.000	0.000
124	A	61	Dark	11.31	-114.97	81.95	0.220	0.000	0.000
125	B	62	Flagged	5.75	9.15	170.85	0.220	0.000	0.000
126	A	62		5.75	9.15	80.85	0.223	0.022	0.002
127	B	63	RGL	4.11	14.71	167.43	0.218	-0.012	0.006
128	A	63	RGL	4.13	14.93	77.43	0.218	-0.012	0.005
129	B	64	RGL	2.61	23.55	160.11	0.214	-0.027	0.017
130	A	64	RGL	2.63	23.66	70.11	0.220	0.003	0.006
131	B	49	Flagged	1.35	49.28	136.17	0.220	0.000	0.000
132	A	65		1.77	-59.61	150.73	0.226	0.024	0.009
133	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
134	B	66	RGL	1.83	-24.49	27.73	0.223	0.058	-0.020
135	A	66	RGL	1.84	-23.68	117.72	0.219	0.006	0.002
136	B	67	RGL	3.28	-13.11	15.42	0.218	-0.012	-0.018
137	A	67	RGL	3.27	-12.75	105.42	0.217	-0.014	0.001
138	B	68	RGL	4.84	-8.81	10.59	0.219	-0.021	-0.003
139	A	68	RGL	4.86	-8.61	100.59	0.218	-0.008	0.001
140	B	69	Dark	6.44	-3.85	8.05	0.190	-0.113	0.024
141	A	69	Dark	6.55	-7.30	98.05	0.288	0.332	0.242
142	B	70	RGL	5.59	-7.70	9.15	0.221	0.006	-0.030
143	A	70	RGL	5.60	-7.56	99.15	0.220	0.017	-0.026

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
144	B	71	RGL	4.03	-10.79	12.57	0.218	-0.024	-0.007
145	A	71	RGL	4.04	-10.52	102.57	0.217	-0.014	0.006
146	B	72	RGL	2.53	-17.38	19.89	0.219	-0.018	-0.014
147	A	72	RGL	2.55	-16.90	109.89	0.217	-0.010	0.009
148	B	73	RGL	1.17	-40.64	43.83	0.226	0.024	-0.019
149	A	73	RGL	1.15	-39.90	133.83	0.230	0.035	0.042
150	B	74	RGL	2.25	-42.66	44.33	0.226	0.073	-0.011
151	A	74	RGL	2.25	-41.83	134.33	0.226	0.069	0.007
152	B	75	RGL	3.56	-25.45	27.15	0.218	0.032	-0.032
153	A	75	RGL	3.57	-25.05	117.15	0.216	0.006	-0.003
154	B	76	RGL	5.02	-17.69	19.17	0.216	0.005	-0.000
155	A	76	RGL	5.03	-17.47	109.17	0.217	0.006	-0.006
156	B	77	RGL	6.53	-13.48	14.73	0.224	0.018	0.019
157	A	77	RGL	6.54	-13.34	104.73	0.223	0.060	-0.004
158	B	78	RGL	5.75	-15.41	16.67	0.220	0.029	0.004
159	A	78	RGL	5.76	-15.23	106.67	0.219	0.022	-0.007
160	B	79	RGL	4.23	-21.19	22.51	0.216	0.025	-0.017
161	A	79	RGL	4.22	-20.71	112.51	0.214	-0.002	0.005
162	B	80	RGL	2.89	-31.91	33.92	0.217	0.007	-0.006
163	A	80	RGL	2.89	-31.41	123.92	0.216	0.011	-0.000
164	B	65	Flagged	1.78	-59.93	60.73	0.220	0.038	-0.013
165	A	81	RGL	3.18	-73.83	163.76	0.220	0.019	0.000
166	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
167	B	82	RGL	2.84	-54.32	55.00	0.218	0.006	-0.015
168	A	82	RGL	2.84	-53.83	145.00	0.218	0.005	-0.005
169	B	83	RGL	3.94	-35.63	36.86	0.217	0.022	-0.020
170	A	83	RGL	3.94	-35.31	126.86	0.216	-0.002	-0.005
171	B	84	RGL	5.30	-25.69	26.94	0.216	0.002	-0.011
172	A	84	RGL	5.30	-25.51	116.94	0.217	0.005	-0.010
173	B	85	RGL	6.76	-19.90	21.03	0.223	0.060	-0.008
174	A	85	RGL	6.77	-19.79	111.03	0.223	0.056	-0.022
175	B	86	RGL	6.00	-22.50	23.64	0.219	0.027	-0.011
176	A	86	RGL	6.00	-22.33	113.64	0.218	0.024	-0.015
177	B	87	RGL	4.61	-30.01	31.21	0.216	0.004	-0.007
178	A	87	RGL	4.62	-29.73	121.21	0.216	0.004	-0.008
179	B	88	RGL	3.36	-43.18	44.52	0.216	0.005	-0.014
180	A	88	RGL	3.36	-42.73	134.52	0.216	0.009	-0.005
181	B	89	RGL	2.45	-69.52	69.03	0.220	0.014	-0.010
182	A	89	RGL	2.44	-68.93	159.03	0.220	0.023	0.001
183	B	90	Flagged	3.71	298.02	61.98	0.247	0.014	0.196
184	A	90		3.71	298.02	151.98	0.217	0.001	-0.002
185	B	91	RGL	4.44	-43.80	44.62	0.217	0.002	0.000
186	A	91	RGL	4.44	-43.56	134.62	0.217	0.013	-0.002
187	B	92	Flagged	5.89	326.22	33.78	0.220	0.000	0.000
188	A	92		5.89	326.22	123.78	0.219	0.002	-0.011
189	B	93	RGL	7.03	-25.87	26.84	0.221	0.050	-0.015
190	A	93	RGL	7.05	-25.76	116.84	0.223	0.061	-0.036
191	B	94	RGL	6.34	-28.98	29.95	0.219	0.025	-0.008
192	A	94	RGL	6.35	-28.80	119.95	0.218	0.017	-0.019
193	B	95	RGL	5.03	-37.64	38.56	0.216	0.004	-0.008
194	A	95	RGL	5.04	-37.43	128.56	0.216	-0.001	-0.008
195	B	96	RGL	3.94	-51.49	52.30	0.216	-0.006	-0.007
196	A	96	RGL	3.94	-51.14	142.30	0.215	-0.006	-0.004
197	B	81	RGL	3.20	-74.19	73.76	0.221	0.018	-0.013

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
198	A	97	RGL	4.67	-78.60	168.87	0.221	0.024	0.008
199	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
200	B	98	RGL	4.20	-66.85	66.78	0.218	0.012	-0.005
201	A	98	RGL	4.20	-66.46	156.78	0.218	0.004	0.002
202	B	99	RGL	4.99	-50.34	50.74	0.216	-0.010	-0.005
203	A	99	RGL	5.01	-49.84	140.74	0.215	-0.003	-0.003
204	B	100	RGL	6.11	-39.09	39.68	0.218	0.014	-0.010
205	A	100	RGL	6.12	-38.92	129.68	0.218	0.006	-0.020
206	B	101	RGL	7.42	-31.37	32.11	0.223	0.043	-0.022
207	A	101	RGL	7.43	-31.26	122.11	0.224	0.046	-0.036
208	B	102	RGL	6.74	-34.76	35.54	0.219	0.026	-0.012
209	A	102	RGL	6.75	-34.61	125.54	0.218	0.020	-0.026
210	B	103	RGL	5.55	-44.11	44.68	0.216	0.003	-0.007
211	A	103	RGL	5.55	-43.91	134.68	0.216	0.001	-0.011
212	B	104	RGL	4.57	-57.65	58.07	0.216	-0.003	-0.009
213	A	104	RGL	4.57	-57.34	148.07	0.216	-0.004	0.001
214	B	105	RGL	3.96	-77.09	76.78	0.221	0.018	-0.007
215	A	105	RGL	3.94	-76.83	166.78	0.219	0.014	0.002
216	B	106	RGL	4.93	-70.09	70.23	0.218	0.007	-0.009
217	A	106	RGL	4.92	-69.85	160.23	0.218	-0.005	-0.003
218	B	107	RGL	5.64	-54.99	55.60	0.215	-0.006	-0.012
219	A	107	RGL	5.65	-54.69	145.60	0.215	-0.022	-0.028
220	B	108	RGL	6.63	-44.22	44.72	0.219	0.016	-0.019
221	A	108	RGL	6.64	-44.04	134.72	0.219	0.003	-0.031
222	B	109	RGL	7.86	-36.10	36.83	0.224	0.034	-0.027
223	A	109	RGL	7.89	-35.94	126.83	0.226	0.045	-0.059
224	B	110	RGL	7.23	-39.80	40.46	0.221	0.021	-0.027
225	A	110	RGL	7.24	-39.66	130.46	0.220	0.012	-0.041
226	B	111	RGL	6.10	-49.15	49.72	0.218	-0.002	-0.010
227	A	111	RGL	6.11	-49.02	139.72	0.216	-0.003	-0.026
228	B	112	RGL	5.22	-62.31	62.43	0.217	0.001	-0.018
229	A	112	RGL	5.22	-62.07	152.43	0.216	-0.012	-0.006
230	B	97	RGL	4.71	-78.87	78.87	0.223	0.024	-0.012
231	A	113		6.22	-82.04	171.55	0.238	0.265	-0.139
232	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
233	B	114	RGL	5.63	-73.03	72.82	0.229	0.063	-0.017
234	A	114	RGL	5.62	-72.68	162.82	0.225	0.035	0.001
235	B	115	RGL	6.29	-59.16	59.50	0.217	-0.003	-0.015
236	A	115	RGL	6.29	-58.97	149.50	0.217	-0.012	-0.015
237	B	116	RGL	7.21	-48.50	49.01	0.221	0.011	-0.024
238	A	116	RGL	7.22	-48.40	139.01	0.221	0.002	-0.040
239	B	117	RGL	8.36	-40.38	41.04	0.229	0.043	-0.054
240	A	117	RGL	8.38	-40.26	131.04	0.231	0.055	-0.084
241	B	118	RGL	7.75	-44.26	44.75	0.224	0.029	-0.039
242	A	118	RGL	7.76	-44.14	134.75	0.223	0.017	-0.059
243	B	119	RGL	6.71	-53.52	53.91	0.217	0.007	-0.019
244	A	119	RGL	6.71	-53.35	143.91	0.217	0.001	-0.030
245	B	120	RGL	6.00	-64.35	65.81	0.203	-0.189	-0.016
246	A	120	RGL	5.92	-65.72	155.81	0.217	-0.008	-0.011
247	B	121	RGL	5.47	-80.79	80.39	0.222	-0.007	0.007
248	A	121	RGL	5.45	-80.61	170.39	0.222	-0.007	-0.001
249	B	122	RGL	6.41	-75.01	74.82	0.223	-0.034	-0.019
250	A	122	RGL	6.40	-74.78	164.82	0.223	-0.033	-0.005
251	B	123	Flagged	7.22	297.34	62.66	0.220	0.000	0.000

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
252	A	123		7.22	297.34	152.66	0.220	-0.045	-0.032
253	B	124		8.07	307.33	52.67	0.226	-0.030	-0.052
254	A	124	Flagged	8.07	307.33	142.67	0.220	0.000	0.000
255	B	125	Dark	8.74	-43.01	44.77	0.233	-0.266	-0.046
256	A	125	Dark	8.77	-44.21	134.77	0.292	0.421	0.119
257	B	126		8.57	311.52	48.48	0.231	-0.010	-0.054
258	A	126	Flagged	8.57	311.52	138.48	0.220	0.000	0.000
259	B	127	RGL	7.36	-57.13	57.39	0.222	-0.041	-0.026
260	A	127	RGL	7.36	-56.93	147.39	0.221	-0.038	-0.047
261	B	128	RGL	6.65	-68.38	68.49	0.218	-0.045	-0.019
262	A	128	RGL	6.64	-68.14	158.49	0.225	-0.025	-0.010
263	B	113	Flagged	6.26	-82.50	81.55	0.165	-0.683	-0.040
264	A	129	RGL	5.47	-100.29	9.61	0.235	0.075	-0.020
265	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
266	B	130	RGL	6.44	-105.81	105.18	0.228	-0.041	0.012
267	A	130	RGL	6.41	-105.83	15.18	0.227	0.010	0.006
268	B	131	RGL	7.05	-118.17	117.34	0.224	-0.052	0.038
269	A	131	RGL	7.01	-118.13	27.34	0.224	-0.053	0.012
270	B	132	RGL	7.94	-128.18	127.33	0.224	-0.074	0.082
271	A	132	RGL	7.89	-128.18	37.33	0.224	-0.048	0.046
272	B	133	Dark	8.19	-135.98	135.23	0.209	-0.482	0.005
273	A	133	Dark	8.65	-132.44	45.23	0.290	0.155	0.005
274	B	134	RGL	8.52	-132.43	131.52	0.259	0.055	0.181
275	A	134	RGL	8.52	-132.65	41.52	0.279	0.192	0.141
276	B	135	RGL	7.45	-123.49	122.61	0.225	-0.058	0.067
277	A	135	RGL	7.42	-123.50	32.61	0.199	-0.045	0.032
278	B	136	RGL	6.69	-112.30	111.51	0.226	-0.033	0.009
279	A	136	RGL	6.67	-112.30	21.51	0.222	-0.023	0.009
280	B	137	RGL	6.25	-98.89	98.45	0.252	0.091	-0.071
281	A	137	RGL	6.24	-98.77	8.45	0.244	0.110	-0.069
282	B	138	RGL	5.66	-107.89	107.19	0.243	0.106	0.002
283	A	138	RGL	5.66	-108.27	17.18	0.241	0.097	0.031
284	B	139	RGL	6.36	-121.64	120.50	0.219	-0.008	0.012
285	A	139	RGL	6.34	-121.69	30.50	0.214	-0.002	0.014
286	B	140	RGL	7.32	-132.01	130.99	0.219	0.015	0.055
287	A	140	RGL	7.30	-131.99	40.99	0.217	0.020	0.038
288	B	141	RGL	8.55	-139.99	138.96	0.248	0.128	0.150
289	A	141	RGL	8.52	-140.09	48.96	0.258	0.210	0.121
290	B	142	RGL	7.79	-135.54	135.25	0.237	0.071	0.099
291	A	142	RGL	7.89	-136.41	45.26	0.239	0.096	0.077
292	B	143	RGL	6.79	-127.02	126.09	0.218	-0.012	0.028
293	A	143	RGL	6.76	-127.06	36.09	0.216	0.000	0.025
294	B	144	RGL	5.99	-115.17	114.19	0.220	-0.000	-0.001
295	A	144	RGL	5.96	-115.15	24.19	0.216	-0.002	-0.000
296	B	129	RGL	5.49	-100.30	99.61	0.234	0.074	-0.017
297	A	145	RGL	3.96	-104.36	13.22	0.220	0.062	0.004
298	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
299	B	146	RGL	4.97	-110.96	109.77	0.219	0.018	-0.005
300	A	146	RGL	4.95	-110.82	19.77	0.219	0.001	0.001
301	B	147	RGL	5.71	-125.77	124.40	0.216	0.013	0.015
302	A	147	RGL	5.70	-125.76	34.40	0.217	-0.004	0.012
303	B	148	RGL	6.74	-136.34	135.28	0.219	0.018	0.036
304	A	148	RGL	6.72	-136.38	45.28	0.219	0.014	0.022
305	B	149	RGL	8.04	-144.18	143.17	0.246	0.153	0.119

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
306	A	149	RGL	8.02	-144.24	53.17	0.252	0.195	0.086
307	B	150	RGL	7.38	-140.67	139.54	0.228	0.065	0.072
308	A	150	RGL	7.35	-140.72	49.54	0.225	0.061	0.044
309	B	151	RGL	6.18	-131.38	130.28	0.218	0.008	0.024
310	A	151	RGL	6.16	-131.44	40.28	0.217	-0.000	0.014
311	B	152	RGL	5.27	-118.71	117.57	0.216	-0.000	-0.004
312	A	152	RGL	5.24	-118.71	27.57	0.216	-0.000	0.004
313	B	153	RGL	4.71	-102.09	101.13	0.222	0.076	-0.013
314	A	153	RGL	4.70	-102.10	11.13	0.230	0.065	0.004
315	B	154	RGL	4.22	-114.54	113.22	0.219	0.015	-0.004
316	A	154	RGL	4.21	-114.47	23.22	0.219	0.005	0.004
317	B	155	RGL	5.10	-130.70	129.26	0.215	-0.003	0.003
318	A	155	RGL	5.08	-130.63	39.26	0.215	0.000	0.004
319	B	156	RGL	6.23	-141.60	140.32	0.217	0.022	0.021
320	A	156	RGL	6.21	-141.68	50.32	0.217	0.019	0.008
321	B	157		7.63	212.11	147.89	0.236	0.118	0.076
322	A	157	Flagged	7.63	212.11	57.89	0.220	0.000	0.000
323	B	158	RGL	6.89	-145.73	144.46	0.225	0.063	0.039
324	A	158	RGL	6.86	-145.77	54.46	0.223	0.050	0.027
325	B	159	RGL	5.63	-136.52	135.32	0.216	0.010	0.016
326	A	159	RGL	5.61	-136.60	45.33	0.213	0.005	0.011
327	B	160	RGL	4.64	-123.68	121.93	0.220	0.057	0.018
328	A	160	RGL	4.62	-123.68	31.93	0.219	0.006	0.005
329	B	145	RGL	3.96	-104.36	103.22	0.226	0.046	-0.019
330	A	161	RGL	2.49	-113.24	20.98	0.230	0.049	0.030
331	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
332	B	162	RGL	3.54	-119.59	118.02	0.222	0.026	-0.002
333	A	162	RGL	3.53	-119.73	28.02	0.221	0.027	0.012
334	B	163		4.66	224.61	135.39	0.216	0.014	0.011
335	A	163	Flagged	4.66	224.61	45.39	0.244	-0.109	0.094
336	B	164	RGL	5.79	-147.53	146.22	0.217	0.017	0.009
337	A	164	RGL	5.77	-147.58	56.22	0.218	-0.004	0.013
338	B	165	Flagged	7.22	-154.25	153.16	0.239	0.150	0.067
339	A	165		7.20	-154.35	63.16	0.236	0.143	0.042
340	B	166	RGL	6.48	-151.25	150.05	0.225	0.068	0.052
341	A	166	RGL	6.45	-151.33	60.05	0.222	0.046	0.017
342	B	167	RGL	5.12	-142.81	141.44	0.216	0.009	0.013
343	A	167	RGL	5.10	-142.88	51.44	0.216	0.001	0.002
344	B	168	RGL	4.01	-129.65	127.70	0.216	0.014	0.006
345	A	168	RGL	3.99	-129.64	37.70	0.217	0.011	0.009
346	B	169	RGL	3.24	-108.09	106.24	0.235	0.018	0.003
347	A	169	RGL	3.23	-107.88	16.24	0.221	0.025	0.012
348	B	170	RGL	2.92	-127.43	125.00	0.216	0.027	0.011
349	A	170	RGL	2.91	-127.46	35.00	0.221	0.017	0.021
350	B	171	RGL	4.05	-144.94	143.14	0.212	0.032	0.013
351	A	171	RGL	4.04	-145.02	53.14	0.217	-0.002	0.015
352	B	172	RGL	5.39	-154.49	153.06	0.215	0.028	0.018
353	A	172	RGL	5.37	-154.60	63.06	0.215	-0.001	0.012
354	B	173	RGL	6.93	-160.08	158.97	0.232	0.118	0.041
355	A	173	RGL	6.91	-160.16	68.97	0.238	0.142	0.017
356	B	174	RGL	6.16	-157.70	156.36	0.223	0.067	0.025
357	A	174	RGL	6.14	-157.76	66.36	0.226	0.051	0.009
358	B	175	RGL	4.72	-150.64	148.79	0.221	0.017	0.005
359	A	175	RGL	4.71	-150.67	58.79	0.217	0.001	0.022

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
360	B	176	RGL	3.44	-137.80	135.48	0.218	0.014	0.011
361	A	176	RGL	3.42	-137.85	45.48	0.215	0.011	0.015
362	B	161	RGL	2.50	-113.01	110.98	0.233	0.061	-0.019
363	A	177	RGL	1.28	-143.98	46.17	0.237	0.103	0.051
364	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
365	B	178	RGL	2.31	-139.01	135.67	0.225	0.037	0.019
366	A	178	RGL	2.31	-138.84	45.67	0.222	0.035	0.029
367	B	179	RGL	3.67	-155.07	152.85	0.217	0.014	0.014
368	A	179	RGL	3.65	-155.22	62.85	0.216	0.006	0.013
369	B	180	RGL	5.14	-162.26	160.83	0.218	0.018	0.012
370	A	180	RGL	5.12	-162.34	70.83	0.218	0.009	-0.003
371	B	181	RGL	6.66	-166.32	165.27	0.233	0.121	0.027
372	A	181	RGL	6.65	-166.45	75.27	0.237	0.138	0.006
373	B	182	RGL	5.91	-164.66	163.33	0.225	0.086	0.019
374	A	182	RGL	5.88	-164.77	73.33	0.224	0.063	-0.002
375	B	183	RGL	4.39	-159.33	157.49	0.216	0.015	0.014
376	A	183	RGL	4.38	-159.42	67.49	0.217	-0.004	0.006
377	B	184	RGL	2.98	-148.93	146.08	0.218	0.016	0.015
378	A	184	RGL	2.95	-149.06	56.08	0.220	0.035	0.010
379	B	185	RGL	1.84	-123.53	119.27	0.237	0.098	0.021
380	A	185	RGL	1.84	-123.42	29.27	0.224	0.087	0.034
381	B	186	RGL	1.94	-156.82	152.28	0.229	0.066	0.013
382	A	186	RGL	1.93	-156.76	62.28	0.225	0.046	0.025
383	B	187	RGL	3.41	-167.23	164.58	0.219	-0.001	0.014
384	A	187	RGL	3.40	-167.28	74.58	0.218	-0.014	0.009
385	B	188	Flagged	4.59	-172.33	169.41	0.220	0.000	0.000
386	A	188		4.87	-170.50	79.41	0.221	-0.013	-0.006
387	B	189	Flagged	6.53	188.05	171.95	0.220	0.000	0.000
388	A	189		6.53	188.05	81.95	0.242	0.131	-0.032
389	B	190	RGL	5.73	-172.08	170.85	0.229	0.054	0.008
390	A	190	RGL	5.71	-172.19	80.85	0.231	0.060	-0.017
391	B	191	RGL	4.17	-169.20	167.43	0.219	-0.005	0.012
392	A	191	RGL	4.16	-169.30	77.43	0.219	-0.029	0.005
393	B	192	RGL	2.63	-163.10	160.11	0.219	0.002	0.014
394	A	192	RGL	2.62	-163.12	70.11	0.220	-0.003	0.015
395	B	177	RGL	1.29	-144.42	136.17	0.243	0.132	0.044
396	A	193	RGL	2.05	118.40	150.73	0.220	0.031	-0.009
397	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
398	B	194		2.04	150.10	27.72	0.223	0.002	-0.018
399	A	194	Flagged	2.06	149.41	117.72	0.223	-0.027	0.013
400	B	195	RGL	3.47	163.09	15.42	0.218	-0.034	-0.022
401	A	195	RGL	3.46	162.79	105.42	0.218	0.006	-0.009
402	B	196	RGL	4.99	168.38	10.59	0.221	-0.007	-0.000
403	A	196	RGL	4.97	168.22	100.59	0.219	-0.001	0.000
404	B	197	RGL	6.49	171.19	8.05	0.225	0.030	0.013
405	A	197	RGL	6.49	171.10	98.05	0.224	0.058	-0.007
406	B	198	RGL	5.75	169.95	9.15	0.223	0.002	0.012
407	A	198	RGL	5.74	169.79	99.15	0.221	0.007	-0.000
408	B	199	RGL	4.23	165.98	12.57	0.219	-0.024	-0.002
409	A	199	RGL	4.22	165.77	102.57	0.218	-0.004	-0.002
410	B	200	RGL	2.74	158.13	19.89	0.219	-0.006	-0.005
411	A	200	RGL	2.73	157.88	109.89	0.220	-0.011	0.005
412	B	201		1.40	125.58	43.83	0.274	-0.017	-0.029
413	A	201	Flagged	6.48	-13.65	133.83	0.220	0.000	0.000

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
414	B	202	RGL	2.53	134.83	44.33	0.238	0.115	-0.028
415	A	202	RGL	2.55	135.09	134.33	0.230	0.087	-0.005
416	B	203	RGL	3.77	151.83	27.15	0.216	0.017	-0.004
417	A	203	RGL	3.77	151.68	117.15	0.216	0.009	-0.006
418	B	204	RGL	5.22	160.03	19.17	0.218	0.010	-0.003
419	A	204	RGL	5.22	159.89	109.17	0.217	0.010	-0.005
420	B	205	RGL	6.69	164.58	14.73	0.221	0.037	0.002
421	A	205	RGL	6.69	164.45	104.73	0.223	0.061	-0.014
422	B	206	RGL	5.93	162.54	16.67	0.219	0.035	-0.002
423	A	206	RGL	5.92	162.40	106.67	0.220	0.034	-0.014
424	B	207	RGL	4.49	156.41	22.51	0.224	0.025	-0.012
425	A	207	RGL	4.48	156.29	112.51	0.225	0.043	-0.014
426	B	208	RGL	3.10	144.48	33.92	0.216	0.006	-0.006
427	A	208	RGL	3.09	144.22	123.92	0.217	0.016	-0.007
428	B	193	RGL	2.06	119.00	60.73	0.225	0.007	-0.017
429	A	209	RGL	3.50	106.56	163.76	0.221	0.019	-0.008
430	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
431	B	210		3.13	125.13	55.00	0.219	0.021	-0.019
432	A	210	Flagged	3.26	125.51	145.00	0.231	-0.065	0.004
433	B	211	RGL	4.20	142.39	36.86	0.215	0.008	-0.010
434	A	211	RGL	4.20	142.27	126.86	0.215	-0.005	-0.007
435	B	212	RGL	5.50	152.29	26.94	0.217	0.019	-0.007
436	A	212	RGL	5.50	152.13	116.94	0.217	0.015	-0.010
437	B	213	RGL	6.96	158.36	21.03	0.222	0.052	-0.006
438	A	213	RGL	6.97	158.24	111.03	0.224	0.062	-0.026
439	B	214	RGL	6.24	155.71	23.64	0.219	0.040	-0.015
440	A	214	RGL	6.23	155.57	113.64	0.218	0.031	-0.017
441	B	215		4.80	148.79	31.21	0.216	0.009	-0.014
442	A	215	Flagged	4.80	148.79	121.21	0.220	0.000	0.000
443	B	216	RGL	3.60	134.48	44.52	0.216	0.007	-0.020
444	A	216	RGL	3.60	134.34	134.52	0.215	0.002	-0.008
445	B	217	RGL	2.78	111.29	69.02	0.222	0.033	-0.011
446	A	217	RGL	2.77	111.11	159.02	0.223	0.037	-0.003
447	B	218	RGL	3.78	118.03	61.98	0.219	0.013	-0.016
448	A	218	RGL	3.79	117.89	151.98	0.218	0.007	-0.008
449	B	219	RGL	4.71	135.02	44.61	0.215	0.009	-0.011
450	A	219	RGL	4.71	134.87	134.61	0.215	0.002	-0.010
451	B	220	RGL	5.90	145.63	33.78	0.216	0.011	-0.006
452	A	220	RGL	5.90	145.51	123.78	0.217	0.005	-0.016
453	B	221	RGL	7.27	152.61	26.84	0.221	0.048	-0.018
454	A	221	RGL	7.27	152.50	116.84	0.223	0.066	-0.040
455	B	222	RGL	6.72	150.20	29.95	0.208	-0.139	-0.016
456	A	222	RGL	6.57	149.34	119.95	0.220	0.022	-0.024
457	B	223	Flagged	5.25	141.44	38.56	0.223	-0.043	0.005
458	A	223		5.25	141.44	128.56	0.215	-0.006	-0.007
459	B	224	RGL	4.20	127.33	52.30	0.217	0.005	-0.017
460	A	224	RGL	4.20	127.13	142.30	0.217	-0.004	-0.010
461	B	209	RGL	3.50	106.72	73.76	0.222	0.000	0.000
462	A	225	OpenSQ	4.94	101.13	168.87	0.220	0.000	0.000
463	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
464	B	226	OpenSQ	4.42	113.22	66.78	0.220	0.000	0.000
465	A	226	OpenSQ	4.42	113.22	156.78	0.220	0.000	0.000
466	B	227	OpenSQ	5.24	129.26	50.74	0.220	0.000	0.000
467	A	227	OpenSQ	5.24	129.26	140.74	0.220	0.000	0.000

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
468	B	228	OpenSQ	6.36	140.32	39.68	0.220	0.000	0.000
469	A	228	OpenSQ	6.36	140.32	129.68	0.220	0.000	0.000
470	B	229	OpenSQ	7.63	147.89	32.11	0.220	0.000	0.000
471	A	229	OpenSQ	7.63	147.89	122.11	0.220	0.000	0.000
472	B	230	OpenSQ	6.98	144.46	35.54	0.220	0.000	0.000
473	A	230	OpenSQ	6.98	144.46	125.54	0.220	0.000	0.000
474	B	231	OpenSQ	5.77	135.32	44.67	0.220	0.000	0.000
475	A	231	OpenSQ	5.77	135.32	134.68	0.220	0.000	0.000
476	B	232	OpenSQ	4.78	121.93	58.07	0.220	0.000	0.000
477	A	232	OpenSQ	4.78	121.93	148.07	0.220	0.000	0.000
478	B	233	OpenSQ	4.17	103.22	76.78	0.220	0.000	0.000
479	A	233	OpenSQ	4.17	103.22	166.78	0.220	0.000	0.000
480	B	234	OpenSQ	5.15	109.77	70.23	0.220	0.000	0.000
481	A	234	OpenSQ	5.15	109.77	160.23	0.220	0.000	0.000
482	B	235	OpenSQ	5.87	124.40	55.60	0.220	0.000	0.000
483	A	235	OpenSQ	5.87	124.40	145.60	0.220	0.000	0.000
484	B	236	OpenSQ	6.88	135.28	44.72	0.220	0.000	0.000
485	A	236	OpenSQ	6.88	135.28	134.72	0.220	0.000	0.000
486	B	237	OpenSQ	8.08	143.17	36.83	0.220	0.000	0.000
487	A	237	OpenSQ	8.08	143.17	126.83	0.220	0.000	0.000
488	B	238	OpenSQ	7.46	139.54	40.46	0.220	0.000	0.000
489	A	238	OpenSQ	7.46	139.54	130.46	0.220	0.000	0.000
490	B	239	OpenSQ	6.35	130.28	49.72	0.220	0.000	0.000
491	A	239	OpenSQ	6.35	130.28	139.72	0.220	0.000	0.000
492	B	240	OpenSQ	5.46	117.57	62.43	0.220	0.000	0.000
493	A	240	OpenSQ	5.46	117.57	152.43	0.220	0.000	0.000
494	B	225	OpenSQ	4.94	101.13	78.87	0.220	0.000	0.000
495	A	241		6.39	99.25	171.55	0.218	0.123	-0.052
496	X	0	DarkSQ	NaN	NaN	NaN	0.000	0.000	0.000
497	B	242		5.92	107.50	72.81	0.220	-0.081	-0.096
498	A	242	Flagged	5.94	107.43	162.81	0.224	0.015	-0.048
499	B	243	RGL	6.57	120.74	59.50	0.223	-0.010	-0.017
500	A	243	RGL	6.58	120.62	149.50	0.218	-0.011	-0.030
501	B	244	RGL	7.49	130.97	49.01	0.221	-0.018	-0.021
502	A	244	RGL	7.50	130.81	139.01	0.220	0.030	-0.068
503	B	245	RGL	8.65	138.94	41.04	0.223	0.058	-0.072
504	A	245	RGL	8.67	138.83	131.04	0.232	0.042	-0.089
505	B	246	RGL	8.04	135.19	44.74	0.222	0.025	-0.053
506	A	246	RGL	8.05	135.05	134.75	0.227	0.006	-0.051
507	B	247		6.97	126.09	53.91	0.218	0.006	-0.030
508	A	247	Flagged	6.97	126.09	143.91	0.220	0.000	0.000
509	B	248	RGL	6.20	114.30	65.81	0.215	0.007	-0.011
510	A	248	RGL	6.21	114.21	155.81	0.216	-0.011	-0.010
511	B	249	RGL	5.76	100.01	80.39	0.219	-0.032	-0.007
512	A	249	RGL	5.76	99.85	170.39	0.220	0.022	-0.003
513	B	250	RGL	6.68	105.39	74.82	0.222	-0.009	-0.009
514	A	250	RGL	6.69	105.30	164.82	0.221	-0.014	-0.007
515	B	251	RGL	7.25	117.40	62.66	0.219	-0.040	-0.028
516	A	251	RGL	7.27	117.29	152.66	0.220	-0.061	-0.040
517	B	252	RGL	8.09	127.23	52.67	0.223	-0.034	0.033
518	A	252	RGL	8.11	127.11	142.67	0.226	-0.046	-0.082
519	B	253	Dark	10.76	132.87	44.77	0.221	0.302	-0.249
520	A	253	Dark	9.70	132.19	134.77	0.236	-0.158	0.141
521	B	254	RGL	8.63	131.46	48.48	0.227	-0.012	-0.064

Continued on next page

GCP ID	Pol	Pair	Notes	r	θ	χ	σ	p	c
522	A	254	RGL	8.65	131.35	138.48	0.224	-0.007	-0.103
523	B	255	RGL	7.65	122.63	57.39	0.220	-0.042	-0.048
524	A	255	RGL	7.67	122.53	147.39	0.225	-0.059	-0.063
525	B	256	RGL	6.93	111.64	68.49	0.214	0.005	-0.014
526	A	256	RGL	6.93	111.44	158.49	0.224	-0.041	-0.021
527	B	241	Flagged	6.40	101.33	81.55	0.228	0.193	0.187

Table D.1: Measured beam properties, polarization orientation angles, and notes for BICEP2 detectors. Detector centroids (r and θ) are derived from an image cross-correlation analysis of CMB maps (as described in Section 3.1.3). The reported polarization orientation angles are derived from the dielectric sheet calibrator measurements, described in Section 3.4.1. The beam width σ, and the beam ellipticity in the 'plus' and 'cross' orientations (p and c, respectively) are derived from far-field beam maps, as reported Section 3.1.3.

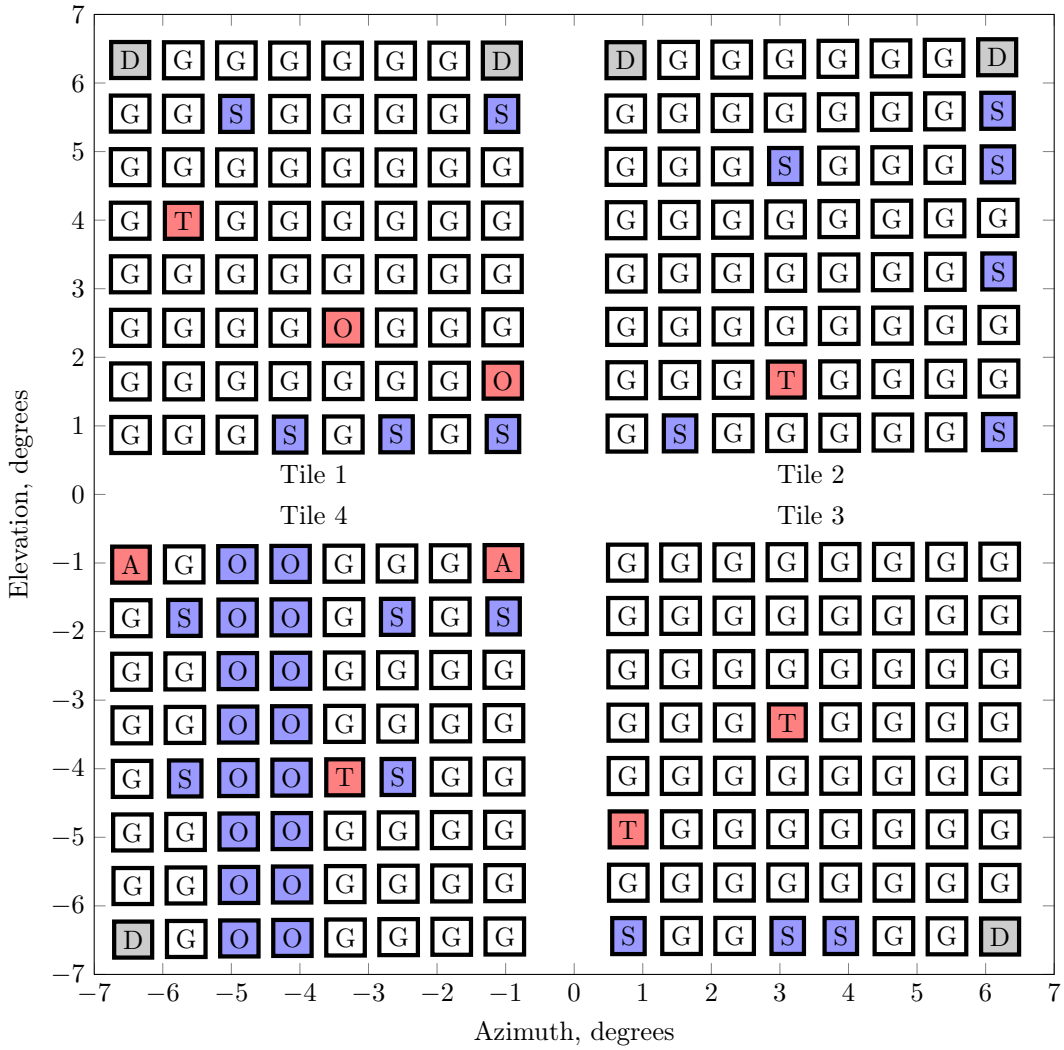


Figure D.1: Guide to the BICEP2 operational focal plane. Total focal plane yield on operational detector pairs is 82%, with the majority of failures coming from open connectivity to one multiplexing column, though not all of these pixels are guaranteed to be included in science data analysis. Blue pixels represent faults at the readout or amplifier stage, red pixels represent faults at the detector stage, and gray represents dark detectors. Letters represent the following:
G: Good operational detector pair
D: Dark detector pair
O: Open - no connectivity to either A or B
S: Non-operational or faulty first-stage SQUID in either A or B
T: Non-operational or faulty TES in either A or B
A: Non-operational or faulty antenna in either A or B

Bibliography

Aikin, R. W., P. A. Ade, S. Benton, J. J. Bock, J. A. Bonetti, J. A. Brevik, C. D. Dowell, L. Duband, J. P. Filippini, S. R. Golwala, M. Halpern, V. V. Hristov, K. Irwin, J. P. Kaufman, B. G. Keating, J. M. Kovac, C. L. Kuo, A. E. Lange, C. B. Netterfield, H. T. Nguyen, R. W. Ogburn, IV, A. Orlando, C. Pryke, S. Richter, J. E. Ruhl, M. C. Runyan, C. Sheehy, S. A. Stokes, R. Sudiwala, G. P. Teply, J. E. Tolan, A. D. Turner, P. Wilson, and C. L. Wong (2010, July). Optical Performance of the BICEP2 Telescope at the South Pole. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 77410V.

Barkats, D. et al. (2013). Improved measurement of cmb polarization power spectra from three years of bicep data. *In Preparation*.

Battistelli, E. S., M. Amiri, B. Burger, M. Halpern, S. Knotek, M. Ellis, X. Gao, D. Kelly, M. Macintosh, K. Irwin, and C. Reintsema (2008, May). Functional Description of Read-out Electronics for Time-Domain Multiplexed Bolometers for Millimeter and Sub-millimeter Astronomy. *Journal of Low Temperature Physics 151*, 908–914.

Bischoff, C. A. (2013). Abscal and Beam Uncertainty for ρ. *BICEP internal note*.

Bischoff, C. A. et al. (2013). ρ from Bandpowers. *BICEP internal note*.

Bond, J. R., A. H. Jaffe, and L. Knox (2000, April). Radical Compression of Cosmic Microwave Background Data. *Astrophys. J. 533*, 19–37.

Bonetti, J. A., P. K. Day, M. Kenyon, C.-L. Kuo, A. Turner, H. G. Leduc, and J. J. Bock (2009, June). Characterization of Antenna-Coupled TES Bolometers for the Spider Experiment. *IEEE Transactions on Applied Superconductivity 19*, 520–523.

Bonetti, J. A., A. D. Turner, J. J. Bock, J. A. Brevik, P. K. Day, J. Filippini, S. R. Golwala, W. Holmes, W. C. Jones, M. Kenyon, J. M. Kovac, C. L. Kuo, H. G. LeDuc, M. Lueker, H. T. Nguyen, R. O'brient, A. Orlando, M. Runyan, Z. Staniszewski, R. Sudiwala, and A. Trangsrud (2012, May). Characterization and Fabrication of the TES Arrays for the Spider, Keck and BICEP2 CMB Polarimeters. *Journal of Low Temperature Physics 167*, 146–151.

Bonetti, J. A., A. D. Turner, M. Kenyon, A. Orlando, J. A. Brevik, A. Trangsrud, R. Sudiwala,

H. G. Leduc, H. T. Nguyen, P. K. Day, J. J. Bock, S. R. Golwala, J. Sayers, J. M. Kovac, A. E. Lange, W. C. Jones, and C. L. Kuo (2009, December). Microfabrication and Device Parameter Testing of the Focal Plane Arrays for the Spider and BICEP2/Keck CMB Polarimeters. In B. Young, B. Cabrera, and A. Miller (Eds.), *American Institute of Physics Conference Series*, Volume 1185 of *American Institute of Physics Conference Series*, pp. 367–370.

Brevik, J. (2012). *Searching for Primordial Gravitational Waves at Degree Scales from the South Pole*. Ph. D. thesis, California Institute of Technology.

Brevik, J. A., R. W. Aikin, M. Amiri, S. J. Benton, J. J. Bock, J. A. Bonetti, B. Burger, C. D. Dowell, L. Duband, J. P. Filippini, S. R. Golwala, M. Halpern, M. Hasselfield, G. Hilton, V. V. Hristov, K. Irwin, J. P. Kaufman, B. G. Keating, J. M. Kovac, C. L. Kuo, A. E. Lange, E. M. Leitch, C. B. Netterfield, H. T. Nguyen, R. W. Ogburn, IV, A. Orlando, C. Pryke, C. Reintsema, S. Richter, J. E. Ruhl, M. Runyan, C. Sheehy, Z. Staniszewski, R. Sudiwala, J. E. Tolan, A. D. Turner, P. Wilson, and C. L. Wong (2010, July). Initial Performance of the BICEP2 Antenna-Coupled Superconducting Bolometers at the South Pole. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 77411H.

Brown, M. L., P. Ade, J. Bock, M. Bowden, G. Cahill, P. G. Castro, S. Church, T. Culverhouse, R. B. Friedman, K. Ganga, W. K. Gear, S. Gupta, J. Hinderks, J. Kovac, A. E. Lange, E. Leitch, S. J. Melhuish, Y. Memari, J. A. Murphy, A. Orlando, C. O'Sullivan, L. Piccirillo, C. Pryke, N. Rajguru, B. Rusholme, R. Schwarz, A. N. Taylor, K. L. Thompson, A. H. Turner, E. Y. S. Wu, M. Zemcov, and QUa D Collaboration (2009, November). Improved Measurements of the Temperature and Polarization of the Cosmic Microwave Background from QUaD. *Astrophys. J. 705*, 978–999.

Cabella, P. and M. Kamionkowski (2004, March). Theory of Cosmic Microwave Background Polarization. *ArXiv Astrophysics e-prints*.

Carroll, S. M. (1998, October). Quintessence and the Rest of the World: Suppressing Long-Range Interactions. *Physical Review Letters 81*, 3067–3070.

Challinor, A. and A. Lewis (2005). Lensed cmb power spectra from all-sky correlation functions. *Phys. Rev. D71*, 103010.

Chervenak, J. A., K. D. Irwin, E. N. Grossman, J. M. Martinis, C. D. Reintsema, and M. E. Huber (1999, June). Superconducting Multiplexer for Arrays of Transition Edge Sensors. *Applied Physics Letters 74*, 4043.

Chiang, C. (2008). *Observation of Cosmic Microwave Background Polarization with BICEP*. Ph. D. thesis, California Institute of Technology.

Chiang, H. C., P. A. R. Ade, D. Barkats, J. O. Battle, E. M. Bierman, J. J. Bock, C. D. Dowell, L. Duband, E. F. Hivon, W. L. Holzapfel, V. V. Hristov, W. C. Jones, B. G. Keating, J. M. Kovac, C. L. Kuo, A. E. Lange, E. M. Leitch, P. V. Mason, T. Matsumura, H. T. Nguyen, N. Ponthieu, C. Pryke, S. Richter, G. Rocha, C. Sheehy, Y. D. Takahashi, J. E. Tolan, and K. W. Yoon (2010, March). Measurement of Cosmic Microwave Background Polarization Power Spectra from Two Years of BICEP Data. *Astrophys. J. 711*, 1123–1140.

Chon, G., A. Challinor, S. Prunet, E. Hivon, and I. Szapudi (2004, May). Fast Estimation of Polarization Power Spectra Using Correlation Functions. *Mon. Not. Roy. Astron. Soc. 350*, 914–926.

de Korte, P. A. J., J. Beyer, S. Deiker, G. C. Hilton, K. D. Irwin, M. Macintosh, S. W. Nam, C. D. Reintsema, L. R. Vale, and M. E. Huber (2003, August). Time-Division Superconducting Quantum Interference Device Multiplexer for Transition-Edge Sensors. *Review of Scientific Instruments 74*, 3807–3815.

Delabrouille, J., M. Betoule, J.-B. Melin, M.-A. Miville-Deschênes, J. Gonzalez-Nuevo, M. Le Jeune, G. Castex, G. de Zotti, S. Basak, M. Ashdown, J. Aumont, C. Baccigalupi, A. Banday, J.-P. Bernard, F. R. Bouchet, D. L. Clements, A. da Silva, C. Dickinson, F. Dodu, K. Dolag, F. Elsner, L. Fauvet, G. Faÿ, G. Giardino, S. Leach, J. Lesgourgues, M. Liguori, J. F. Macias-Perez, M. Massardi, S. Matarrese, P. Mazzotta, L. Montier, S. Mottet, R. Paladini, B. Partridge, R. Piffaretti, G. Prezeau, S. Prunet, S. Ricciardi, M. Roman, B. Schaefer, and L. Toffolatti (2012, July). The Pre-Launch Planck Sky Model: A Model of Sky Emission at Submillimetre to Centimetre Wavelengths. *ArXiv e-prints*.

Dodelson, S. (2003). *Modern Cosmology*. Academic Press, Waltham, MA.

Dunkley, J., E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, J. L. Weiland, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright (2009, February). Five-Year Wilkinson Microwave Anisotropy Probe Observations: Likelihoods and Parameters from the WMAP Data. *Astrophys. J. Suppl. 180*, 306–329.

Feldman, G. J. and R. D. Cousins (1998, April). Unified Approach to the Classical Statistical Analysis of Small Signals. *Phys. Rev. D 57*, 3873–3889.

Finkbeiner, D. P., M. Davis, and D. J. Schlegel (1999, October). Extrapolation of Galactic Dust Emission at 100 Microns to Cosmic Microwave Background Radiation Frequencies Using FIRAS. *Astrophys. J. 524*, 867–886.

Gear, W. K. and C. R. Cunningham (1990, December). SCUBA: A Submillimetre Camera. In B. Kaldeich (Ed.), *Liege International Astrophysical Colloquia*, Volume 29 of *Liege International Astrophysical Colloquia*, pp. 353–358.

Hinshaw, G., D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright (2012, December). Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *ArXiv e-prints*.

Hivon, E., K. M. Górski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen (2002, March). MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. *Astrophys. J. 567*, 2–17.

Hu, W. and S. Dodelson (2002). Cosmic Microwave Background Anisotropies. *ArXiv Astrophysics e-prints*.

Hu, W., M. M. Hedman, and M. Zaldarriaga (2003, February). Benchmark Parameters for CMB Polarization Experiments. *Phys. Rev. D 67*(4), 043004.

Irwin, K. D. and G. C. Hilton (2005). Transition-Edge Sensors. *Cryogenic Particle Detection*, 63.

Jones, W. C., T. E. Montroy, B. P. Crill, C. R. Contaldi, T. S. Kisner, A. E. Lange, C. J. MacTavish, C. B. Netterfield, and J. E. Ruhl (2007, August). Instrumental and Analytic Methods for Bolometric Polarimetry. *Astr. & Astroph. 470*, 771–785.

Kaufman, J. P. et al. (2013a, February). Thermal Filter Transfer Function. *BICEP2 internal note*.

Kaufman, J. P. et al. (2013b). Thermal Stability of the BICEP2 Telescope. *In preparation*.

Kernasovskiy, S., P. A. R. Ade, R. W. Aikin, M. Amiri, S. Benton, C. Bischoff, J. J. Bock, J. A. Bonetti, J. A. Brevik, B. Burger, G. Davis, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, M. Halpern, M. Hasselfield, G. Hilton, V. V. Hristov, K. Irwin, J. M. Kovac, C. L. Kuo, E. Leitch, M. Lueker, C. B. Netterfield, H. T. Nguyen, R. O'Brient, R. W. Ogburn, C. L. Pryke, C. Reintsema, J. E. Ruhl, M. C. Runyan, R. Schwarz, C. D. Sheehy, Z. Staniszewski, R. Sudiwala, G. Teply, J. E. Tolan, A. D. Turner, A. Vieregg, D. V. Wiebe, P. Wilson, and C. L. Wong (2012, September). Optimization and Sensitivity of the Keck Array. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 8452 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 84521B.

Knox, L. (1999, November). Cosmic Microwave Background Anisotropy Window Functions Revisited. *Phys. Rev. D 60*(10), 103516.

Lamb, J. W. (1996, December). Miscellaneous Data on Materials for Millimetre and Submillimetre Optics. *International Journal of Infrared and Millimeter Waves 17*, 1997–2034.

Liddle, A. R. and D. H. Lyth (2000). *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press.

Lide, D. R. (2007, June). *CRC Handbook of Chemistry and Physics, 88th Edition*. CRC Press, Boca Raton, FL.

Little, A. and D. Lyth (2000). *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, Cambridge.

Miville-Deschênes, M.-A., N. Ysard, A. Lavabre, N. Ponthieu, J. F. Macías-Pérez, J. Aumont, and J. P. Bernard (2008, November). Separation of Anomalous and Synchrotron Emissions Using WMAP Polarization Data. *Astr. & Astroph. 490*, 1093–1102.

O'Brient, R., P. A. R. Ade, Z. Ahmed, R. W. Aikin, M. Amiri, S. Benton, C. Bischoff, J. J. Bock, J. A. Bonetti, J. A. Brevik, B. Burger, G. Davis, P. Day, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, J. Grayson, M. Halpern, M. Hasselfield, G. Hilton, V. V. Hristov, H. Hui, K. Irwin, S. Kernasovskiy, J. M. Kovac, C. L. Kuo, E. Leitch, M. Lueker, K. Megerian, L. Moncelsi, C. B. Netterfield, H. T. Nguyen, R. W. Ogburn, C. L. Pryke, C. Reintsema, J. E. Ruhl, M. C. Runyan, R. Schwarz, C. D. Sheehy, Z. Staniszewski, R. Sudiwala, G. Teply, J. E. Tolan, A. D. Turner, R. S. Tucker, A. Vieregg, D. V. Wiebe, P. Wilson, C. L. Wong, W. L. K. Wu, and K. W. Yoon (2012, September). Antenna-coupled TES bolometers for the Keck array, Spider, and Polar-1. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 8452 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 84521G.

O'Dea, D., A. Challinor, and B. R. Johnson (2007, April). Systematic Errors in Cosmic Microwave Background Polarization Measurements. *Mon. Not. Roy. Astron. Soc. 376*, 1767–1783.

O'Dell, C. (2002, January). A New Upper Limit on the Polarization of the Cosmic Microwave Background Radiation. *ArXiv Astrophysics e-prints*.

Ogburn, R. W., P. A. R. Ade, R. W. Aikin, M. Amiri, S. J. Benton, C. A. Bischoff, J. J. Bock, J. A. Bonetti, J. A. Brevik, E. Bullock, B. Burger, G. Davis, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, M. Gordon, M. Halpern, M. Hasselfield, G. Hilton, V. V. Hristov, H. Hui, K. Irwin, J. P. Kaufman, B. G. Keating, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, E. M. Leitch, M. Lueker, T. Montroy, C. B. Netterfield, H. T. Nguyen, R. O'Brient, A. Orlando, C. L. Pryke, C. Reintsema, S. Richter, J. E. Ruhl, M. C. Runyan, R. Schwarz, C. D. Sheehy, Z. K. Staniszewski, R. V. Sudiwala, G. P. Teply, K. Thompson, J. E. Tolan, A. D. Turner, A. G. Vieregg, D. V. Wiebe, P. Wilson, and C. L. Wong (2012, September). BICEP2 and Keck Array Operational Overview and Status of Observations. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 8452 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 84521A.

Ogburn, R. W., P. A. R. Ade, R. W. Aikin, M. Amiri, S. J. Benton, J. J. Bock, J. A. Bonetti, J. A. Brevik, B. Burger, C. D. Dowell, L. Duband, J. P. Filippini, S. R. Golwala, M. Halpern, M. Hasselfield, G. Hilton, V. V. Hristov, K. Irwin, J. P. Kaufman, B. G. Keating, J. M. Kovac, C. L. Kuo, A. E. Lange, E. M. Leitch, C. B. Netterfield, H. T. Nguyen, A. Orlando, C. L. Pryke, C. Reintsema, S. Richter, J. E. Ruhl, M. C. Runyan, C. D. Sheehy, Z. K. Staniszewski, S. A. Stokes, R. V. Sudiwala, G. P. Teply, J. E. Tolan, A. D. Turner, P. Wilson, and C. L. Wong (2010, July). The BICEP2 CMB Polarization Experiment. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 77411G.

Penzias, A. A. and R. W. Wilson (1965, July). A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophys. J. 142*, 419–421.

Pryke, C., P. Ade, J. J. Bock, M. Bowden, M. L. Brown, G. Cahill, P. G. Castro, S. Church, T. Culverhouse, R. Friedman, K. Ganga, W. K. Gear, S. Gupta, J. Hinderks, J. J. Kovac, A. E. Lange, E. Leitch, S. J. Melhuish, Y. Memari, J. A. Murphy, A. Orlando, R. Schwarz, C. O'Sullivan, L. Piccirillo, N. Rajguru, B. Rusholme, A. N. Taylor, K. L. Thompson, A. H. Turner, E. Y. S. Wu, and M. Zemcov (2008). *Astrophys. J.*, submitted, (astro–ph/0805.1944).

Reichardt, C. L., P. A. R. Ade, J. J. Bock, J. R. Bond, J. A. Brevik, C. R. Contaldi, M. D. Daub, J. T. Dempsey, J. H. Goldstein, W. L. Holzapfel, C. L. Kuo, A. E. Lange, M. Lueker, M. Newcomb, J. B. Peterson, J. Ruhl, M. C. Runyan, and Z. Staniszewski (2009, April). High-Resolution CMB Power Spectrum from the Complete ACBAR Data Set. *Astrophys. J. 694*, 1200–1219.

Reichardt, C. L., L. Shaw, O. Zahn, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, J. Dudley, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, M. Joy, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, T. Natoli, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, and R. Williamson (2012, August). A Measurement of Secondary Cosmic Microwave Background Anisotropies with Two Years of South Pole Telescope Observations. *Astrophys. J. 755*, 70.

Runyan, M. C. and W. C. Jones (2008, September). Thermal Conductivity of Thermally-Isolating Polymeric and Composite Structural Support Materials Between 0.3 and 4 K. *Cryogenics 48*, 448–454.

Seljak, U. and M. Zaldarriaga (1997, March). Signature of Gravity Waves in the Polarization of the Microwave Background. *Physical Review Letters 78*, 2054–2057.

Shimon, M., B. Keating, N. Ponthieu, and E. Hivon (2008, April). CMB Polarization Systematics Due to Beam Asymmetry: Impact on Inflationary Science. *Phys. Rev. D 77*(8), 083003.

Smith, K. M. (2006, October). Pseudo-C_ℓ Estimators which do not Mix E and B-modes. *Phys. Rev. D 74*(8), 083002.

Stiehl, G. M., H. M. Cho, G. C. Hilton, K. D. Irwin, J. A. B. Mates, C. D. Reintsema, and B. L. Zink (2011, June). Time-Division SQUID Multiplexers With Reduced Sensitivity to External Magnetic Fields. *IEEE Transactions on Applied Superconductivity 21*, 298–301.

Story, K. T., C. L. Reichardt, Z. Hou, R. Keisler, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, J. Dudley, B. Follin, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, J. D. Hrubes, M. Joy, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, J. T. Sayre, K. K. Schaffer, L. Shaw, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, A. van Engelen, K. Vanderlinde, J. D. Vieira, R. Williamson, and O. Zahn (2012, October). A Measurement of the Cosmic Microwave Background Damping Tail from the 2500-square-degree SPT-SZ survey. *ArXiv e-prints*.

Takahashi, Y. (2010). *Measurement of the Cosmic Microwave Background Polarization with the Bicep Telescope at the South Pole*. Ph. D. thesis, University of California, Berkeley.

Takahashi, Y. D., P. A. R. Ade, D. Barkats, J. O. Battle, E. M. Bierman, J. J. Bock, H. C. Chiang, C. D. Dowell, L. Duband, E. F. Hivon, W. L. Holzapfel, V. V. Hristov, B. G. Keating, J. M. Kovac, C. L. Kuo, A. E. Lange, E. M. Leitch, P. V. Mason, H. T. Nguyen, N. Ponthieu, C. Pryke, S. Richter, G. D. Rocha, and K. W. Yoon (2010, March). Characterization of the bicep telescope for high-precision cosmic microwave background polarimetry. *Astrophys. J. 711*, 1141–1157.

Tolan, J. E. (2013). Synchrotron and Dust Foregrounds. *BICEP internal note*.

White, M., J. E. Carlstrom, M. Dragovan, and W. L. Holzapfel (1999, March). Interferometric Observation of Cosmic Microwave Background Anisotropies. *Astrophys. J. 514*, 12–24.

Wright, A. E., M. R. Griffith, B. F. Burke, and R. D. Ekers (1996, October). The Parkes-MIT-NRAO 4.85GHz (PMN) Surveys (Griffith+ 1993-1996). *VizieR Online Data Catalog 8038*, 0.

Wright, E. L., X. Chen, N. Odegard, C. L. Bennett, R. S. Hill, G. Hinshaw, N. Jarosik, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, J. L. Weiland, E. Wollack, J. Dunkley, B. Gold, M. Halpern, A. Kogut, D. Larson, M. Limon, S. S. Meyer, and G. S. Tucker (2009, February). Five-Year Wilkinson Microwave Anisotropy Probe Observations: Source Catalog. *Astrophys.*

J. Suppl. 180, 283–295.

Yadav, A. P. S., M. Su, and M. Zaldarriaga (2010, March). Primordial B-mode Diagnostics and Self-Calibrating the CMB Polarization. *Phys. Rev. D 81*(6), 063512.

Yoon, K. (2007). *Design and Deployment of BICEP: A Novel Small-Aperture CMB Polarimeter to Test Inflationary Cosmology*. Ph. D. thesis, California Institute of Technology.