BICEP2 / Keck Array V: Measurements of B-mode Polarization at Degree Angular Scales and 150 GHz by the Keck Array

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:33717523

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

BICEP2/KECK ARRAY. VII. MATRIX BASED *E/B* SEPARATION APPLIED TO BICEP2 AND THE KECK ARRAY

P. A. R. Ade[1], Z. Ahmed[2,3], R. W. Aikin[4], K. D. Alexander[5], D. Barkats[5], S. J. Benton[6], C. A. Bischoff[5], J. J. Bock[4,7], R. Bowens-Rubin[5], J. A. Brevik[4], I. Buder[5], E. Bullock[8], V. Buza[5,9], J. Connors[5], B. P. Crill[7], L. Duband[10], C. Dvorkin[9], J. P. Filippini[4,11], S. Fliescher[8], J. Grayson[2], M. Halpern[12], S. Harrison[5], S. R. Hildebrandt[4,7], G. C. Hilton[13], H. Hui[4], K. D. Irwin[2,3], J. Kang[2,3], K. S. Karkare[5], E. Karpel[2], J. P. Kaufman[14], B. G. Keating[14], S. Kefeli[4], S. A. Kernasovskiy[1], J. M. Kovac[5,9], C. L. Kuo[2,3], E. M. Leitch[15], M. Lueker[4], K. G. Megerian[7], T. Namikawa[2,3], C. B. Netterfield[6,16], H. T. Nguyen[7], R. O'Brient[4,7], R. W. Ogburn IV[2,3], A. Orlando[4], C. Pryke[8,17], S. Richter[5], R. Schwarz[17], C. D. Sheehy[15,17], Z. K. Staniszewski[4,7], B. Steinbach[4], R. V. Sudiwala[1], G. P. Teply[4,14], K. L. Thompson[2,3], J. E. Tolan[2], C. Tucker[1], A. D. Turner[7], A. G. Vieregg[5,15,18], A. C. Weber[7], D. V. Wiebe[12], J. Willmert[17], C. L. Wong[5,9], W. L. K. Wu[2,19], AND K. W. Yoon[2]
(KECK ARRAY AND BICEP2 COLLABORATIONS)

[1] School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, UK
[2] Department of Physics, Stanford University, Stanford, CA 94305, USA; jetolan@stanford.edu
[3] Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, Menlo Park, CA 94025, USA
[4] Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA
[5] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street MS 42, Cambridge, MA 02138, USA
[6] Department of Physics, University of Toronto, Toronto, Ontario, M5S 1A7, Canada
[7] Jet Propulsion Laboratory, Pasadena, CA 91109, USA
[8] Minnesota Institute for Astrophysics, University of Minnesota, Minneapolis, MN 55455, USA
[9] Department of Physics, Harvard University, Cambridge, MA 02138, USA
[10] Service des Basses Températures, Commissariat à l'Energie Atomique, F-38054 Grenoble, France
[11] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[12] Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia, V6T 1Z1, Canada
[13] National Institute of Standards and Technology, Boulder, CO 80305, USA
[14] Department of Physics, University of California at San Diego, La Jolla, CA 92093, USA
[15] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[16] Canadian Institute for Advanced Research, Toronto, Ontario, M5G 1Z8, Canada
[17] School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
[18] Department of Physics, Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA
[19] Department of Physics, University of California, Berkeley, CA 94720, USA

ABSTRACT

A linear polarization field on the sphere can be uniquely decomposed into an *E*-mode and a *B*-mode component. These two components are analytically defined in terms of spin-2 spherical harmonics. Maps that contain filtered modes on a partial sky can also be decomposed into *E*-mode and *B*-mode components. However, the lack of full sky information prevents orthogonally separating these components using spherical harmonics. In this paper, we present a technique for decomposing an incomplete map into *E* and *B*-mode components using *E* and *B* eigenmodes of the pixel covariance in the observed map. This method is found to orthogonally define *E* and *B* in the presence of both partial sky coverage and spatial filtering. This method has been applied to the BICEP2 and the Keck Array maps and results in reducing *E* to *B* leakage from ΛCDM *E*-modes to a level corresponding to a tensor-to-scalar ratio of $r < 1 \times 10^{-4}$.

Key words: cosmic background radiation – cosmology: observations – gravitational waves – inflation – polarization

1. INTRODUCTION

Current experiments are producing low noise maps of the polarization of the cosmic microwave background (CMB) radiation able to constrain models of inflation and measure *B*-modes from gravitational lensing. These experiments include BICEP2, the Keck Array, POLARBEAR, SPTPOL, ACTPol, and *Planck* (Hanson et al. 2013; BICEP2 Collaboration I 2014; Polarbear Collaboration 2014; Keck Array & BICEP2 Collaborations VI 2015; Planck Collaboration I 2015; van Engelen et al. 2015). These experiments do not measure the CMB over the entire sky for a variety of reasons. Galactic foregrounds prevent any experiment from producing a map of the CMB over the entire sky. Any ground or balloon based experiment has a limited view of the full sky. Some

experiments, including BICEP2 and the Keck Array, choose to observe a limited field of view to increase map depth over a small region of sky or choose to filter their data so that the maps incompletely measure the modes within the field.

The ability to uniquely separate a linear polarization field into *E* and *B*-modes is critical for measuring gravitational waves using the *B*-mode polarization. This separation allows the distinction to be made between *E*-modes created by scalar perturbations and *B*-modes coming from tensor perturbations (Kamionkowski et al. 1997; Zaldarriaga 1998).

Unfortunately, the unique decomposition into *E* and *B* is only possible for maps of the full sky. Maps containing a limited view of the sky, or an incomplete measurement of the true sky modes, are said to suffer from *E/B* leakage. *E* to *B*

leakage is defined as measured power for a particular B-mode estimator whose source is true sky E-mode power. B to E leakage is leakage of power in the opposite direction, but in practice it is less of a concern for CMB measurements due to the much fainter B-mode signal. E/B leakage refers to both types of leakage.

There are several ways to mitigate the effect of E/B leakage in analysis. Full pixel-space likelihood methods in principle can optimally separate E and B contributions for any given map. These have been applied mainly to maps of relatively modest pixel count, including many early detections of CMB polarization (for example, Kovac et al. 2002; Readhead et al. 2004; Bischoff et al. 2008). Current analyses more commonly apply fixed estimators of E and B power spectra to observed CMB polarization maps. The simplest way to correct such estimators for leakage is to run an ensemble of simulations through the analysis and subtract the mean level of leakage in the angular power spectrum. However, the sample variance from the leaked power remains and contributes to the final uncertainty of measured power in each angular power spectrum bin, limiting an experiment's ability to measure B-modes regardless of its instrumental sensitivity. For many experiments, including BICEP2 and the Keck Array, the sample variance of the leaked E-modes is comparable to the instrumental noise and is a significant contribution to the uncertainty in the B-mode power spectrum.

Solutions to this problem rely on the fact that for most B-mode science it is not necessary to classify all the modes in the measured polarization field. Instead, it is sufficient to find subspaces that are caused by either E or B and ignore the modes whose source cannot be determined. There are a number of published methods that attempt this goal.

Smith (2006) presents an estimator that does not suffer from E/B leakage arising from partial sky coverage. This method has been incorporated into the xpure and S^{2}HAT packages (Grain et al. 2009), and the BICEP2 and Keck Array analysis pipeline contains an option in which this algorithm is implemented.

However, many experiments, including BICEP2 and the Keck Array, produce maps in which some modes have also been removed by filtering. The estimator presented in Smith (2006) does not prevent filtered modes from creating E/B leakage. Another method, presented in Smith & Zaldarriaga (2007), accounts for incomplete mode measurement in partial sky maps. However, we have found this method to be computationally infeasible for the BICEP2 and Keck Array observing and filtering strategy.

For BICEP2 and the Keck Array, we developed a new method for distinguishing true sky B-mode polarization from the leaked E-modes in the observed maps. The method extends the work of Bunn et al. (2003) and applies it to a real data set. It is a standard component of the BICEP2 and Keck Array analysis pipeline and effectively eliminates the uncertainty created by E/B leakage. The method reduces the final uncertainty in the measured BB power spectrum of the BICEP2 results (BICEP2 Collaboration I 2014) by more than a factor of two, compared to analysis done with the Smith (2006) method. The method results in a larger improvement for the analysis of the combined BICEP2 and Keck Array maps (Keck Array & BICEP2 Collaborations VI 2015), where the noise levels are lower.

The organization of this paper is as follows: Section 2 provides an abbreviated background of a polarization field on a

sphere, decomposition into spin-2 spherical harmonics, and analytically defines E and B-modes. Section 3 outlines the eigenvalue problem used in the matrix based E/B separation. Section 4 describes how an observation matrix is created in the BICEP2 and Keck Array analysis pipeline. Section 5 describes constructing the signal covariance matrix and Section 6 uses the covariance matrix to solve the eigenvalue problem and find purification matrices. Section 7 presents results of matrix based E/B separation in the BICEP2 data set. Concluding remarks are offered in Section 8.

Unless otherwise stated, we adopt the HEALPix polarization convention[20] and work in J2000.0 equatorial coordinates throughout this paper. Bold font letters and symbols represent vectors or matrices, even when containing subscripts, in which case the subscript is meant to designate a new matrix or vector. Normal font letters and symbols represent scalar quantities.

2. E AND B-MODES FROM A POLARIZATION FIELD

This section demonstrates the decomposition of a polarization field on the full sky into E and B-modes. Much of the discussion follows Zaldarriaga & Seljak (1997) and Bunn et al. (2003).

2.1. Full Sky

The values of the Stokes parameters Q and U for a particular location on the sky are dependent on the choice of coordinate system. By rotating the local coordinate system, Q is rotated into U and vice versa. Under rotation by an angle ϕ, the combinations $Q + iU$ and $Q - iU$ transform as:

$$\begin{aligned}(Q + iU)' &= e^{-2i\phi}(Q + iU) \\ (Q - iU)' &= e^{2i\phi}(Q - iU).\end{aligned} \quad (1)$$

The T, Q, and U fields can be expressed as sums of spin weighted spherical harmonics. While the temperature anisotropies can be broken down into spin-0 harmonics, the polarization field of Q and U must be expressed in terms of spin-2 spherical harmonics (Goldberg et al. 1967):

$$\begin{aligned}T(\boldsymbol{r}) &= \sum_{lm} a_{lm}^T \, (_0 Y_{lm}(\boldsymbol{r})) \\ (Q + iU)(\boldsymbol{r}) &= \sum_{lm} a_{+2,lm} (_{+2} Y_{lm}(\boldsymbol{r})) \\ (Q - iU)(\boldsymbol{r}) &= \sum_{lm} a_{-2,lm} (_{-2} Y_{lm}(\boldsymbol{r})),\end{aligned} \quad (2)$$

where $_{\pm 2} Y_{lm}$ are the spin-2 case of spin weighted spherical harmonics, and the spin-0 case are the normal spherical harmonics, $_0 Y_{lm}$. Since $Q + iU$ and $Q - iU$ are affected by rotations of the coordinate system, it is convenient to express the coefficients of the spin-2 spherical harmonics using a set of coordinate independent scalar a_{lm}^E coefficients and pseudoscalar a_{lm}^B coefficients:

$$\begin{aligned}a_{lm}^E &\equiv -(a_{+2,lm} + a_{-2,lm})/2 \\ a_{lm}^B &\equiv -i(a_{+2,lm} - a_{-2,lm})/2.\end{aligned} \quad (3)$$

We also define two combinations of spin-2 spherical harmonics:

$$X_{1,lm} \equiv (_{+2}Y_{lm} + {}_{-2}Y_{lm})/2$$
$$X_{2,lm} \equiv (_{+2}Y_{lm} - {}_{-2}Y_{lm})/2. \qquad (4)$$

We can use the coefficients in Equation (3) and the combinations in Equation (4) to construct real space forms of T, Q, and U fields, according to Equation (2):

$$T(\mathbf{r}) = \sum_{lm} a_{lm}^T ({}_0Y_{lm}(\mathbf{r}))$$
$$Q(\mathbf{r}) = - \sum_{lm} (a_{lm}^E X_{1,lm}(\mathbf{r}) + i a_{lm}^B X_{2,lm}(\mathbf{r}))$$
$$U(\mathbf{r}) = - \sum_{lm} (a_{lm}^B X_{1,lm}(\mathbf{r}) - i a_{lm}^E X_{2,lm}(\mathbf{r})). \qquad (5)$$

Using these relations, we can write the polarization field as a vector:

$$P(\mathbf{r}) \equiv \begin{pmatrix} Q(\mathbf{r}) \\ U(\mathbf{r}) \end{pmatrix}$$
$$= - \sum_{lm} \begin{bmatrix} a_{lm}^E X_{1,lm}(\mathbf{r}) + i a_{lm}^B X_{2,lm}(\mathbf{r}) \\ a_{lm}^B X_{1,lm}(\mathbf{r}) - i a_{lm}^E X_{2,lm}(\mathbf{r}) \end{bmatrix}$$
$$= - \sum_{lm} \begin{bmatrix} a_{lm}^E \begin{pmatrix} X_{1,lm}(\mathbf{r}) \\ -i X_{2,lm}(\mathbf{r}) \end{pmatrix} + a_{lm}^B \begin{pmatrix} i X_{2,lm}(\mathbf{r}) \\ X_{1,lm}(\mathbf{r}) \end{pmatrix} \end{bmatrix}$$
$$= - \sum_{lm} [a_{lm}^E Y_{lm}^E(\mathbf{r}) + a_{lm}^B Y_{lm}^B(\mathbf{r})], \qquad (6)$$

where Y_{lm}^E and Y_{lm}^B have been introduced and defined in the last step. On the full sphere, Y_{lm}^E and Y_{lm}^B are orthogonal:

$$\int_{S^2} Y_{lm}^E(\mathbf{r}) \cdot Y_{l'm'}^B(\mathbf{r}) dS = 0, \qquad (7)$$

for all l, l' and m, m'.

2.2. Orthogonality of Pure E and Pure B

The inner product of two polarization fields is defined as:

$$P \cdot P' \equiv \int_{\Omega} P \cdot P' d\Omega, \qquad (8)$$

where Ω is the manifold on which the polarization field is defined: for the full sky it is the celestial sphere. In pixelized maps, the vector space of a polarization field has a finite dimension: twice the number of pixels in the map.

As demonstrated in Equation (7), E and B-mode polarization fields on the full sky are orthogonal. However, experiments produce Q/U maps of portions of the sky, and often filter spatial modes out of these maps. We define the term "observed" maps or modes to refer to these incomplete measurements of the true sky.

The spaces of observed E-modes and B-modes are non-orthogonal. The overlapping subspace between the two is called the ambiguous space. We cannot tell whether signal in the ambiguous subspace came from full sky E-modes or full sky B-modes.

The solution is to decompose vector fields on an observed manifold into three subspaces: "pure" E-modes, "pure" B-modes, and ambiguous modes. Pure E and B-modes are subspaces of the polarization vector space of a particular manifold, defined as:

1. A pure B-mode is orthogonal to observed E-modes.
2. A pure E-mode is orthogonal to observed B-modes.

Therefore, a pure B-mode is one that has no E to B leakage: neither pure E-modes nor ambiguous modes contribute to it.

3. HOW MATRIX BASED E/B SEPARATION FINDS PURE E AND PURE B

A pure B-mode on an observed manifold is defined in Section 2.2 as being orthogonal to observed E-modes:

$$P^E \cdot b = 0. \qquad (9)$$

The vector b is any linear combination of modes in the subspace of the pure B-modes. For pixelized maps, b contains Q and U values for each of the pixels in the map, and P^E is the pixelized version of the E-mode spherical harmonics. It is useful to multiply the above equation by its conjugate transpose, and sum over l and m, so that we have a scalar representing the degree of orthogonality:

$$b^{\top} \left(\sum_{lm} a_{lm}^{E*} a_{lm}^E Y_{lm}^E Y_{lm}^{E\dagger} \right) b = 0. \qquad (10)$$

We have freedom to choose the power spectrum, $C_l^{EE} = \langle a_{lm}^{E*} a_{lm}^E \rangle$, which is included in the covariance matrix, C_E:

$$C_E \equiv \sum_{lm} C_l^{EE} Y_{lm}^E Y_{lm}^{E\dagger}. \qquad (11)$$

We note that this product is the 2×2 $[Q, U]$ covariance block in the signal covariance matrix:

$$C_E = \langle P^E (P^E)^{\top} \rangle$$
$$= \begin{pmatrix} \langle Q_i^E Q_j^E \rangle & \langle Q_i^E U_j^E \rangle \\ \langle U_i^E Q_j^E \rangle & \langle U_i^E U_j^E \rangle \end{pmatrix}, \qquad (12)$$

where the superscript denotes the E-mode component of the full sky polarization field and i, j designate pixels in the map. We can evaluate the covariance matrix for a particular set of pixels and a chosen spectrum.

By solving a generalized eigenvalue equation of the form:

$$C_B x_i = \lambda_i C_E x_i, \qquad (13)$$

and selecting eigenmodes corresponding to the largest eigenvalues, we can find eigenmodes b that are nearly orthogonal to E-modes and therefore approximate pure B. Eigenmodes corresponding to the smallest eigenvalues approximate pure E. This method is a natural extension to the signal to noise truncation discussed in Bond et al. (1998) and Bunn & White (1997) and applied in Kuo et al. (2004). The specific application to E and B-modes was first discussed in Bunn et al. (2003).

We say the modes approximate pure E and pure B-modes because the degree of orthogonality is proportional to the magnitude of the eigenvalues. The level of orthogonality is discussed further in Section 6. However, for the remainder of the paper, we will use the terms pure B and pure E to refer to the largest and smallest eigenmodes of Equation (13), despite the fact that their inner product is not identically zero.

Now suppose that the true sky polarization field, P, is transformed into an observed polarization field, $\tilde{P}$, by a real space linear operation, R:

$$\tilde{P}^E = RP^E$$
$$= -R\sum_{lm}a_{lm}^E Y_{lm}^E. \tag{14}$$

Throughout this paper, transformations into observed quantities are indicated by the inclusion of a tilde over the variable, in the above equation, $P \rightarrow \tilde{P}$. The operator R will typically represent filtering operations necessary to suppress noise and/or systematics plus an apodization of the resulting observed maps.

The condition for pure E and pure B must be the same after multiplying by R. We still demand that the vectors of pure B be orthogonal to all those in the E space, which includes both the pure E-modes and the ambiguous modes:

$$\left(R\sum_{lm}a_{lm}^E Y_{lm}^E\right)\cdot b = 0. \tag{15}$$

We create a basis of pure E and pure B-modes by solving the eigenvalue problem with the covariances of the form:

$$\tilde{C}_E = R^\top\left(\sum_{lm}C_l^{EE}Y_{lm}^E Y_{lm}^{E\dagger}\right)R$$
$$\tilde{C}_B = R^\top\left(\sum_{lm}C_l^{BB}Y_{lm}^B Y_{lm}^{B\dagger}\right)R, \tag{16}$$

so that Equation (13) becomes:

$$\tilde{C}_B x_i = \lambda_i \tilde{C}_E x_i. \tag{17}$$

In the simplest case, the matrix R is an apodization window and filled only on its diagonal. However, Equation (15) does not necessitate that the real space operator be a diagonal matrix. Any analysis steps that can be expressed as linear operations can be included.

In BICEP2 and Keck Array analysis a number of filtering operations are typically performed during the map making process. In the next section the matrix R corresponding to these operations is derived. The practical implementation of a solution to the eigenvalue equation is discussed in Section 6.

4. OBSERVATION MATRIX

The matrix R transforms an "input map," m, a vector of the true sky polarization field, into a vector of the observed map, $\tilde{m}$. If the matrix R represents the apodization and linear filtering of an analysis pipeline, it is defined to be the "observation" matrix for a particular experiment. This choice of R ensures the eigenspaces of Equation (17) are pure E and B for the observed map. This section describes how the observation matrix is computed for BICEP2 and the Keck Array.

The steps in constructing the observation matrix mirror functions in the data reduction pipeline that was originally developed for QUAD (Pryke et al. 2009) and later used in the BICEP1 (BICEP1 Collaboration 2014), BICEP2 (BICEP2 Collaboration I 2014), and Keck Array (Keck Array & BICEP2 Collaborations V 2015) analyses. This pipeline consists of a MATLAB library of procedures which constructs maps, including several filtering steps, from real data or simulated timestream data for a given input sky map.

The filtering operations performed sequentially in the standard pipeline include data selection, polynomial filtering, scan-synchronous signal subtraction, weighting, binning into map pixels, and deprojection of leaked temperature signal. To construct the observation matrix, matrices representing each of these steps are multiplied together to form a final matrix that performs all of the operations at once. Since each of the operations is linear, the observation matrix is independent of the input map. Therefore, the same matrix can be used on any input map and will perform the same operations as the standard pipeline.

If the combined matrix of timestream operations is $\mathcal{V}$, then transforming a timestream, d, into an observed map, $\tilde{m}$, is simply:

$$\tilde{m} = \mathcal{V}d. \tag{18}$$

The signal component of a timestream can be generated from an input map, m, using a matrix that contains information about the pointing and orientations of the detectors, according to the equation $d = \mathcal{A}m$. The observation matrix, R, is given by the product of $\mathcal{V}$ and $\mathcal{A}$:

$$\tilde{m} = \mathcal{V}\mathcal{A}m \tag{19}$$
$$= Rm. \tag{20}$$

It is not necessary for the input maps and observed maps to share the same pixelization scheme, since the observation matrix can easily be made to transform between the two.

4.1. Input HEALPix Maps

We choose a HEALPix pixelization scheme (Gorski et al. 2005) for the input maps, m, because it has equal area pixels on the sphere and is widely used in the cosmology community.

A true sky signal is represented by the map $m^o = [T_{xy}^o, Q_{xy}^o, U_{xy}^o]$, where (x, y) are the (R.A., decl.) coordinates of the map. Using synfast[21], the unobserved input map is convolved with the array averaged beam function, $\bar{\mathcal{B}}$, constructed from measurements of the beam function of all detectors in the array:

$$\begin{bmatrix} T_{xy} \\ Q_{xy} \\ U_{xy} \end{bmatrix} = \bar{\mathcal{B}} * m^o = \bar{\mathcal{B}}_{xy} * \begin{bmatrix} T_{xy}^o \\ Q_{xy}^o \\ U_{xy}^o \end{bmatrix}. \tag{21}$$

The input map vector is found by reforming the beam convolved two-dimensional map into a one-dimensional vector, m, of length $3j$, where $j = 1 \ldots n_p$, for n_p pixels in the input map:

$$m \equiv \begin{bmatrix} T_j \\ Q_j \\ U_j \end{bmatrix}. \tag{22}$$

4.2. BICEP2 and Keck Array Scan Strategy

The observing strategies for BICEP2 and the Keck Array are very similar and borrow heavily from BICEP1. All three experiments target a region of sky centered at a right ascension

---

[21] synfast is a program in the HEALPix suite that renders sky maps from sets of input a_{lm}'s.

of $0°$ and declination of $-57°.5$. A detailed description of the scan strategy is contained in BICEP2 Collaboration II (2014).

1. *Halfscans:* During normal observations, the telescope scans in azimuth at a constant elevation. The scan speed of $2°.8\,s^{-1}$ in azimuth places the targeted multipoles of $20 < l < 200$ at temporal frequencies less than 1 Hz. Each scan covers $64°.2$ in azimuth, at the end of which the telescope stops and reverses direction in azimuth and scans back across the field center. A scan in a single direction is known as a "halfscan."

2. *Scansets:* Halfscans are grouped into sets of ≈ 100 halfscans, which are known as "scansets." The scan pattern deliberately covers a fixed range in azimuth within each scanset, rather than a fixed range in right ascension. Over the course of the 50 minute scanset, Earth's rotation results in a relative drift of azimuthal coordinates and right ascension of about $12°.5$. At the end of each scanset the elevation is offset by $0°.25$, and a new scanset commences. The telescope steps in $0°.25$ elevation increments between each scanset. All observations take place at 20 elevation steps, with a boresight pointing ranging in elevation between $55°$ and $59°.75$. The geographic location of the telescope, near the South Pole, means that elevation and declination are approximately interchangeable.

3. *Phases:* Scansets are grouped together into sets known as "phases." For BICEP2 and the Keck Array, CMB phases consist of ten scansets, comprising 9 hours of observations. CMB phases are grouped into seven types and each type has a unique combination of elevation offset and azimuthal position.

4. *Schedules:* The third degree of freedom in the BICEP2 and Keck Array telescope mounts is a rotation about the boresight, referred to as "deck rotation." The polarization angles relative to the cryostats are fixed, so rotating in deck angle allows detector pairs to observe at multiple polarization angles.

     A "schedule" typically consists of a set of phases at a particular deck angle. The deck angle is rotated between schedules. There is typically one schedule per fridge cycle, occurring every ~ 3 days for BICEP2 and ~ 2 days for the Keck Array.

4.3. Relationship Between Timestreams and [T, Q, U]

The BICEP2 and Keck Array detectors consist of pairs of nominally co-pointed, orthogonal, polarization sensitive phased array antennas coupled to TES bolometers (BICEP2/Keck & Spider Collaborations 2015). The signal in the timestream from detector "A" is:

$$\tau_t^A = T_t + \cos(2\Psi_t^A)Q_t + \sin(2\Psi_t^A)U_t, \qquad (23)$$

where $[T_t, Q_t, U_t]$ are the Stokes parameters of the beam convolved sky signal for timestream sample t. A timestream consists of n_t time ordered measurements of the sky, $t = 1 \ldots n_t$. Ψ^A is the angle the "A" antenna makes with the Q, U axis on the sky. For the HEALPix polarization convention, this axis is a vector pointing toward the north celestial pole.

The relative gain normalized "A" timestream is summed and differenced with the normalized timestream from its orthogonal
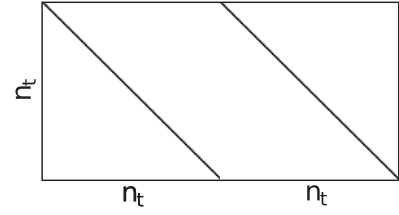

Figure 1. Detector orientation matrix, $[\alpha^- \ \beta^-]$. The matrix is only filled on the diagonals of the two sub-blocks, α and β.

partner "B":

$$s_t = \frac{1}{2}(\tau_t^A + \tau_t^B) = T_t + \alpha_t^+ Q_t + \beta_t^+ U_t$$
$$d_t = \frac{1}{2}(\tau_t^A - \tau_t^B) = \alpha_t^- Q_t + \beta_t^- U_t. \qquad (24)$$

The variables α and β are defined by:

$$\alpha_t^\pm \equiv \frac{1}{2}[\cos(2\Psi_t^A) \pm \cos(2\Psi_t^B)]$$
$$\beta_t^\pm \equiv \frac{1}{2}[\sin(2\Psi_t^A) \pm \sin(2\Psi_t^B)], \qquad (25)$$

where Ψ^B is the angle the "B" antenna makes with the Q, U axis on the sky. Assuming that "A" and "B" are perfectly co-pointed and orthogonal, the signal portion of the timestream vectors can be described with a transformation, A_{tj}, from the input map pixel (with index j) to the timestream sample (with index t):

$$s_t = A_{tj}T_j$$
$$d_t = \alpha_t^- A_{tj}Q_j + \beta_t^- A_{tj}U_j, \qquad (26)$$

where the terms α^+ and β^+ in the pair sum timestream cancel due to the orthogonal orientation of the "A" and "B" detectors. The signal-plus-noise timestreams, in vector notation, are:

$$s = \frac{1}{2}(n_A + n_B) + A[T] \qquad (27)$$

$$d = \frac{1}{2}(n_A - n_B) + [\alpha^- \ \beta^-]\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}\begin{bmatrix} Q \\ U \end{bmatrix}. \qquad (28)$$

where n_A and n_B are the time ordered noise components of detector "A" and detector "B," assuming the noise is uncorrelated with the pointing of the detector pair. For signal only simulations, n_A and n_B can be ignored.

The matrix $[\alpha^- \ \beta^-]$ contains the information about the orientation of a pair's antennas relative to Q and U defined on the sky. We call it the detector orientation matrix. The combination:

$$[\alpha^- \ \beta^-]\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \qquad (29)$$

transforms input Q, U maps into a pair difference timestream. $[\alpha^- \ \beta^-]$ is constructed from two diagonal matrices, α^- and β^-, which are filled with the sine and cosine of the detector orientations at each time sample. A graphical representation of the detector orientation matrix is shown in Figure 1. (Additional steps accounting for polarization efficiency and pair non-orthogonality are absorbed into a normalization correction to the pair difference timestream.)
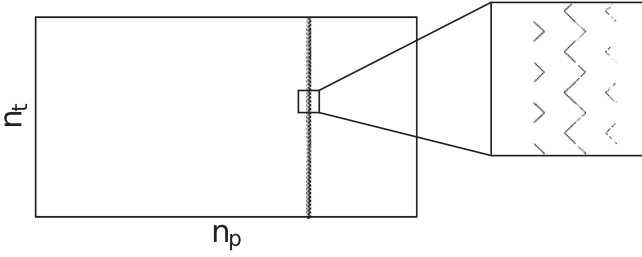
Figure 2. Timestream forming matrix, A: filled elements of the matrix that takes HEALPix maps to timestreams. This matrix contains the pointing of a single detector pair over one scanset within a $N_{\mathrm{side}} = 512$ HEALPix map. The pattern of the filled elements is determined by the particular HEALPix pixel indexing scheme. There are n_t filled entries, consisting of a 1 for each timestream sample. Note that although the above image appears to have multiple pointing locations for a single timestream sample, n_t, this is merely a result of limited resolution in the image. The timestream forming matrix contains only one HEALPix pixel location for each time sample.

4.4. Timestream Forming Matrix, A

The matrix $A = A_{tj}$ represents the timestream forming matrix for a detector pair. It transforms the input temperature map, T_j, into the signal component of the pair sum timestream, s_t. A graphical representation of the timestream forming matrix is shown in Figure 2.

To create timestreams with smooth transitions at pixel boundary crossings, the input maps should have a resolution higher than the spatial band limit imposed by the beam function. For this reason, $N_{\mathrm{side}} = 512$ HEALPix maps are used, whose pixels have a Nyquist frequency $\gtrsim 2\times$ the band limit of the BICEP2 and Keck Array 150 GHz beam function.

The current BICEP2 and Keck Array CMB observations fall within the region of sky bounded in right ascension by $-3^{\mathrm{h}}40^{\mathrm{m}} < \alpha < 3^{\mathrm{h}}40^{\mathrm{m}}$ and in declination by $-70° < \delta < -45°$. This region contains $n_p = 111{,}593$ pixels in an $N_{\mathrm{side}} = 512$ HEALPix map. The number of samples in a scanset is typically $n_t \approx 43{,}000$.

The simplest form of A performs nearest neighbor interpolation of the HEALPix maps, in which case A is $(n_t \times n_p)$ and is filled with ones where the detector pair is pointed and zeros otherwise.

A more sophisticated form of A performs Taylor interpolation on the HEALPix map, in which case A is $\left(n_t \times \frac{\lambda(\lambda+1)}{2} n_p\right)$, where λ is the order of the Taylor polynomial used in interpolation. In this case, A is a matrix that performs Taylor interpolation, allowing sub-pixel accuracy to be recovered from the input map, and m must also contain derivatives of the true sky temperature and polarization field. This matrix is used to build the deprojection templates in Section 4.10 but is not used for forming timestreams because it increases the dimensions of the observation matrix, making the computation of the observation matrix more difficult.

4.5. Polynomial Filtering Matrix, F

To remove low frequency atmospheric noise from the data, a third order polynomial is fit and subtracted from each halfscan in the timestreams. Since each halfscan traces an approximately constant elevation trajectory across the target field, the polynomial filter removes power only in the right ascension direction of the maps. In multipole, l, the third order polynomial filter typically rolls off power below $l < 40$. This can be represented by a "filtering matrix," F, which is block
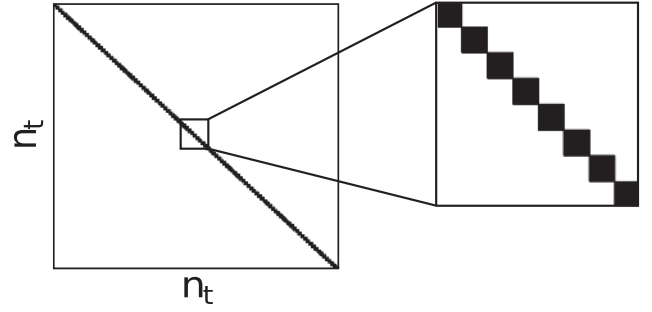

Figure 3. Polynomial filtering matrix, F, showing the filled elements of the matrix. The matrix is very sparse and is block diagonal with blocks the size of a halfscan (≈ 404 samples).

diagonal with the block size being the temporal length of a halfscan. Each block is composed of a matrix:

$$F = I - V(V^{\top}V)^{-1}V^{\top}, \tag{30}$$

where I is the identity matrix and V is the same third order Vandermonde matrix for each halfscan of equal length. The Vandermonde matrix is defined as:

$$V_{tj} = x_t^{j-1}, \tag{31}$$

where $j = 4$ for a third order filter and x_t are the coordinate locations.

For BICEP2 and the Keck Array, x_t is a vector of the relative azimuthal location of each sample in the halfscan. A representation of the polynomial filtering matrix is shown in Figure 3.

4.6. Scan-synchronous Signal Removal Matrix, G

Scan-synchronous subtraction removes signal in the timestreams that is fixed relative to the ground rather than moving with the sky. These azimuthally fixed signals are decoupled from signals rotating with the sky by the scan strategy, which observes over a fixed range in azimuth as the sky slides by (as described in Section 4.2). A template of the mean azimuthal signal is subtracted from the timestreams for each scan direction. This procedure can be represented as a matrix operator, referred to as a "scan-synchronous signal matrix."

The mean azimuthal signal can be found using a matrix $X = X_{tt'}$, for which each row is only filled for entries containing the same azimuthal pointing as the diagonal entry. The scan-synchronous signal matrix subtracts off the mean azimuthal signal:

$$G = I - X, \tag{32}$$

where I is the identity matrix. A graphical representation of the scan-synchronous signal removal matrix is shown in Figure 4. Note that while the F matrix is block diagonal and sparse, and the G matrix is sparse, once the two are combined, the resulting filter matrix is neither sparse nor block diagonal, making matrix operations more computationally demanding.

4.7. Inverse Variance Weighting Matrices, $w^{\pm}$

The timestreams are weighted based on the measured inverse variance of each scanset. Pair sum and pair difference are weighted separately from weights calculated from the two timestreams, w^+ and w^-. The scheme assigns lower weight to particularly noisy channels and periods of bad weather. This
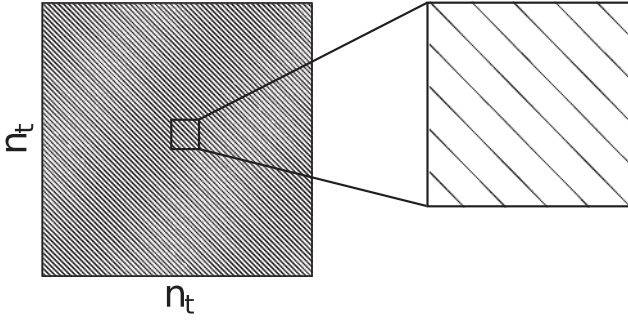
Figure 4. Scan-synchronous signal removal matrix, G, showing the filled elements. The scan-synchronous signal matrix is sparse Toeplitz, with off diagonal components that subtract the average scan-synchronous signal for one of the two scan directions in a scanset.

choice of weighting is not a fully "optimal" map maker (Tegmark 1997), but instead represents a practical solution that avoids calculating and inverting a large noise covariance matrix. The weighting is represented by a matrix whose diagonal is filled with the vector $w^+ = w_{tt}^+$ for pair sum and $w^- = w_{tt}^-$ for pair difference, shown in Figure 5.

4.8. Filtered Signal Timestream Generation

Ignoring noise and combining all the operators of Sections 4.5, 4.6 and 4.7, the sum and difference timestreams in Equation (26) are tranformed to the filtered timestreams:

$$\tilde{s} = w^+ GFA[T]$$

$$\tilde{d} = w^- GF \begin{bmatrix} \alpha^- & \beta^- \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} Q \\ U \end{bmatrix}, \quad (33)$$

the second of which is graphically represented in Figure 6.

4.9. Pointing Matrix, Λ

The timestream quantities s and d are converted to maps by the pointing matrix, $\Lambda = \Lambda_{it}$. If the pixelization of the input maps were identical to the output maps, the pointing matrix would be the transpose of the timestream forming matrix:

$$\Lambda = A^\top.$$

As discussed in Section 4.1, the input maps are HEALPix $N_{\text{side}} = 512$. However, the BICEP maps instead use a simple rectangular grid of pixels in R.A. and decl.: the size of the pixels is $0°25$ in decl., with the pixel size in R.A. set to be equivalent to $0°25$ of arc at the mid-declination of the map, resulting in $236 \times 100 = 23{,}600$ pixels. If the BICEP maps are naively used as "flat maps" then projection distortions are inherent. However, note that the A and Λ matrices together fully encode the mapping from the underlying curved sky to the observed map pixels, allowing such distortions to be accounted for. Figure 2 shows A for a single detector over a scanset and Figure 7 shows Λ for a single detector over a scanset.

The pointing matrix for a single detector pair can be used to construct a pair sum "pairmap:"

$$m_T = \Lambda s = \Lambda w^+ GFA[T]. \quad (34)$$

The pair difference timestream is converted into pairmaps using two copies of the pointing matrix. The two pair difference

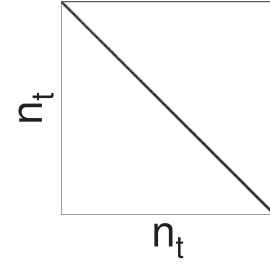**Figure 5.** Weighting matrices $w^\pm$, showing the filled elements. The weighting matrices are zero except on the diagonal, where they contain the weights based on the inverse variance of the timestream.
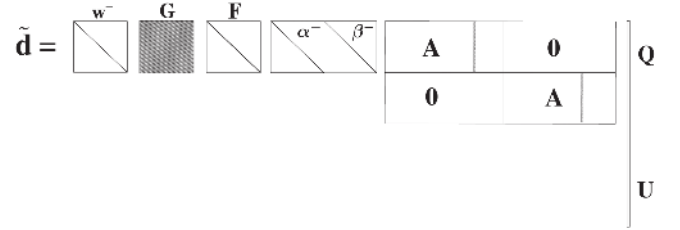

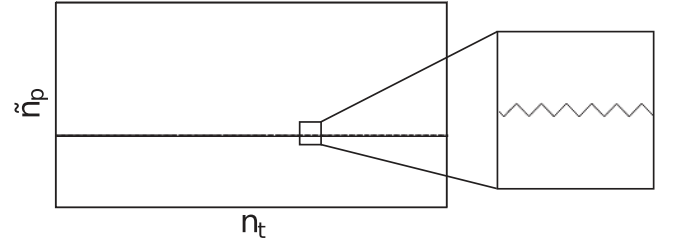**Figure 6.** Matrix generation of simulated timestreams corresponding to Equation (33).

Figure 7. Pointing matrix, Λ: filled elements of the pointing matrix that transforms timestreams to an observed map in the BICEP pixelization. This matrix contains the mapping between the pointing of a single detector pair over one scanset and the output map pixels. There are 23,600 pixels in a BICEP map, denoted as $\tilde{n}_p$. There are n_t filled entries, consisting of a 1 for each timestream sample. Each leg of the zigzag pattern corresponds to a halfscan within the scanset, where the telescope is scanning back and forth at a fixed elevation.

pairmaps correspond to linear combinations of Stokes Q and U:

$$\begin{bmatrix} m_{\alpha^-} \\ m_{\beta^-} \end{bmatrix} = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \begin{bmatrix} \alpha^- \\ \beta^- \end{bmatrix} w^- GF \begin{bmatrix} \alpha^- & \beta^- \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} Q \\ U \end{bmatrix}. \quad (35)$$

For later convenience in abbreviating this equation, we define:

$$\mathcal{P} \equiv \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \begin{bmatrix} \alpha^- \\ \beta^- \end{bmatrix} w^- GF \begin{bmatrix} \alpha^- & \beta^- \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}. \quad (36)$$

4.10. Deprojection Matrix, D

A potential systematic concerning polarization measurements is the leakage of unpolarized signal into polarized signal. In the case of CMB polarization, this takes the form of the relatively bright temperature anisotropy leaking into the much fainter polarization anisotropy. The leakage is caused by imperfect differencing between the orthogonal pairs of detectors. The beam functions can be well approximated by elliptical Gaussians, the difference of which correspond to gain, pointing, width and ellipticity (Hu et al. 2003; Shimon et al. 2008).

The Bicep2 and Keck Array pipeline removes leaked temperature signal from the polarization signal using linear regression to fit leakage templates to the polarization data. This method allows the beam mismatch parameters to be fitted directly from the CMB data itself, rather than relying on external calibration data sets, and is robust to temporal variations of the beam mismatch.

The templates used in the regression are constructed from *Planck* 143 GHz temperature maps.[22] These maps contain both CMB and foreground emission at approximately the Bicep2 band. The noise in *Planck* 143 GHz is significantly subdominant to the CMB temperature anisotropy. For a full description and derivation of the deprojection technique, see Aikin (2013), Sheehy (2013), BICEP1 Collaboration (2014), and BICEP2 Collaboration III (2015). In this section, the entire deprojection algorithm is re-cast as a matrix operation.

For the purposes of generating deprojection templates, we use a timestream forming matrix that performs Taylor expansion around the nearest pixel center to the detector pointing location. The Taylor interpolating matrix produces higher fidelity timestreams than a nearest neighbor matrix. This is important for the deprojection algorithm since small displacements in beam position are responsible for the systematic effect that is removed. Without Taylor interpolation, pixel boundary discontinuities introduce noise and limit the effectiveness of deprojection.

A Taylor polynomial of order λ has $\frac{\lambda(\lambda+1)}{2}$ terms, so the dimensions of the input map vector for second order interpolation is $1 \times 6n_p$. Using Equation (21), the input maps are convolved with the array averaged beam function. The smoothing is done using `synfast`, which contains the ability to output derivatives of the temperature (and polarization) field. Because the beam is applied first, the output derivatives are less noisy than they would be in the raw maps.

The maps are of the form:

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{T} \\ \boldsymbol{\nabla}_\theta \boldsymbol{T} \\ \boldsymbol{\nabla}_\phi \boldsymbol{T} \\ \boldsymbol{\nabla}_{\theta\theta} \boldsymbol{T} \\ \boldsymbol{\nabla}_{\phi\phi} \boldsymbol{T} \\ \boldsymbol{\nabla}_{\theta\phi} \boldsymbol{T} \end{bmatrix}, \tag{37}$$

where θ and ϕ are the `HEALPix` map's latitude and longitude.

Using the temperature map and its derivatives, we can find the Taylor interpolated temperature timestream by replacing $\boldsymbol{A}$ with a Taylor interpolating matrix, $\boldsymbol{A}'$:

$$\boldsymbol{A}' = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{A}\boldsymbol{\Delta\theta} & \boldsymbol{A}\boldsymbol{\Delta\phi} & \boldsymbol{A}\frac{\Delta\theta^2}{2} & \boldsymbol{A}\frac{\Delta\phi^2}{2} & \boldsymbol{A}\boldsymbol{\Delta\theta}\boldsymbol{\Delta\phi} \end{bmatrix}, \tag{38}$$

where $\boldsymbol{\Delta\theta}$ and $\boldsymbol{\Delta\phi}$ are diagonal matrices giving the difference between the detector pair's pointing and the nearest `HEALPix` pixel center.

A differential beam generating operator is applied to the timestreams to create differential beam timestreams. For example, the differential gain timestream is just the beam convolved temperature field:

$$\boldsymbol{d}_{\delta g} = \delta_g \boldsymbol{A}' \boldsymbol{\Theta}, \tag{39}$$

where the fit coefficient for the gain mismatch is δ_g. The differential pointing components are found from the first derivatives of the temperature field with respect to the focal plane coordinates, x and y:

$$\boldsymbol{d}_{\delta x} = \delta_x \, \boldsymbol{\nabla}_x \boldsymbol{A}' \boldsymbol{\Theta} \tag{40}$$

$$\boldsymbol{d}_{\delta y} = \delta_y \, \boldsymbol{\nabla}_y \boldsymbol{A}' \boldsymbol{\Theta}, \tag{41}$$

where δ_x and δ_y are the differential beam coefficients and $\boldsymbol{\nabla}_x$ and $\boldsymbol{\nabla}_y$ are partial differential operators with respect to the focal plane coordinates. Further details of this calculation and derivations for other beam modes are discussed in Appendix C of Bicep2 Collaboration III (2015).

The differential beam timestreams are transformed into maps analogously to Equation (35), creating a pairmap template, $\tilde{\mathcal{T}}_j$, for each differential beam mode, j. The template pairmaps for each scanset, $\mathbb{S}$, are then coadded over phases. For instance, the template pairmap for differential gain, is $\tilde{\mathcal{T}}_1$:

$$\tilde{\mathcal{T}}_1 = \sum_{\mathbb{S}}^{\mathbb{S}\in\text{phase}} \left(\begin{bmatrix} \boldsymbol{\Lambda} & 0 \\ 0 & \boldsymbol{\Lambda} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}^- \\ \boldsymbol{\beta}^- \end{bmatrix} \boldsymbol{w}^- \boldsymbol{GF} \right.$$
$$\left. \times \begin{bmatrix} \boldsymbol{\alpha}^- & \boldsymbol{\beta}^- \end{bmatrix} \begin{bmatrix} \boldsymbol{A}' & 0 \\ 0 & \boldsymbol{A}' \end{bmatrix} \begin{bmatrix} \boldsymbol{\Theta} \\ \boldsymbol{\Theta} \end{bmatrix} \right)_{\mathbb{S}}. \tag{42}$$

A matrix performing weighted linear least-squares regression against real pairmaps produces the fitted coefficients for each of the differential beam modes, $\boldsymbol{c} \equiv [\delta_g, \delta_x, \delta_y, \ldots]$:

$$\boldsymbol{c} = (\tilde{\mathcal{T}}^\top (\boldsymbol{\mathcal{W}}^-)^{-1} \tilde{\mathcal{T}})^{-1} \tilde{\mathcal{T}}^\top (\boldsymbol{\mathcal{W}}^-)^{-1} \begin{bmatrix} \boldsymbol{m}_{\alpha^-} \\ \boldsymbol{m}_{\beta^-} \end{bmatrix}, \tag{43}$$

where $\begin{bmatrix} \boldsymbol{m}_{\alpha^-} \\ \boldsymbol{m}_{\beta^-} \end{bmatrix}$ is the real data pairmap coadded over a phase, $\tilde{\mathcal{T}}$ is a vector of pairmap templates, and $\boldsymbol{\mathcal{W}}^-$ is the pair difference weight map, created from the weight matrix according to:

$$\boldsymbol{\mathcal{W}}^- = \sum_{\mathbb{S}}^{\mathbb{S}\in\text{phase}} \left(\begin{bmatrix} \boldsymbol{\Lambda}\boldsymbol{\alpha}^-\boldsymbol{\alpha}^-\boldsymbol{w}^-\boldsymbol{\Lambda}^\top \\ \boldsymbol{\Lambda}\boldsymbol{\beta}^-\boldsymbol{\beta}^-\boldsymbol{w}^-\boldsymbol{\Lambda}^\top \end{bmatrix} \right)_{\mathbb{S}}. \tag{44}$$

The pairmap templates weighted by $\boldsymbol{c}$ are then subtracted from the real data pairmap:

$$\begin{bmatrix} \tilde{\boldsymbol{m}}_{\alpha^-} \\ \tilde{\boldsymbol{m}}_{\beta^-} \end{bmatrix} = \begin{bmatrix} \boldsymbol{m}_{\alpha^-} \\ \boldsymbol{m}_{\beta^-} \end{bmatrix} - \tilde{\mathcal{T}}\boldsymbol{c}. \tag{45}$$

This process takes the form of a matrix operator that includes each of the beam systematics, giving the deprojection matrix:

$$\boldsymbol{D} \equiv \boldsymbol{I} - \tilde{\mathcal{T}} (\tilde{\mathcal{T}}^\top (\boldsymbol{\mathcal{W}}^-)^{-1} \tilde{\mathcal{T}})^{-1} \tilde{\mathcal{T}}^\top (\boldsymbol{\mathcal{W}}^-)^{-1}. \tag{46}$$

Deprojected pairmaps are then found according to:

$$\begin{bmatrix} \tilde{\boldsymbol{m}}_{\alpha^-} \\ \tilde{\boldsymbol{m}}_{\beta^-} \end{bmatrix} = \boldsymbol{D} \sum_{\mathbb{S}}^{\mathbb{S}\in phase} \mathcal{P}_{\mathbb{S}} \begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{U} \end{bmatrix}. \tag{47}$$

The regression in Equation (43) operates simultaneously over all the modes to be deprojected. Because the templates for different modes are not in general orthogonal, the coefficient for each mode depends on the full set of modes. Therefore, the subtraction in Equation (46) must include the same mode list used in the regression in Equation (43). If the regression included more modes than the subtraction step, the regression would have extra degrees of freedom. This could result in incomplete removal of leakage signal. We avoid this possibility

---

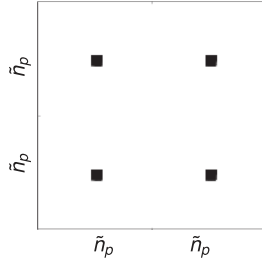[22] For the Keck Array 95 and 220 GHz bands, we use *Planck* 100 GHz and 217 GHz maps.

Figure 8. Deprojection matrix, $\boldsymbol{D}$: filled elements of the deprojection matrix for one pair, for one phase of data. The overall dimensions are $2\tilde{n}_p \times 2\tilde{n}_p$, twice the number of pixels in a BICEP map.

by deferring the regression step until immediately before the subtraction step, and explicitly using the same mode selection for both.

As in the standard pipeline, we deproject for each detector pair, after coadding scanset to phases. To reduce the computational demands, the matrix deprojection pairmaps have additionally been coadded over scan direction, whereas the standard pipeline performs regression separately for left going and right going scans. This is the only difference between simulations run with the standard pipeline and those calculated from the observation matrix and leads to a negligible difference, see Figures 11 and 12.

The deprojection matrix made for a phase is less sparse than one made for a scanset because over the course of a phase a particular pair will observe a larger range of elevation than it would in a scanset. The filled elements of the matrix $\boldsymbol{D}$, for one pair across one phase, is shown in Figure 8.

4.11. Coadding Over Scansets and Detector Pairs to form the Observation Matrix

An observed temperature map, $\tilde{\boldsymbol{T}}'$, can be found by summing the pair sum pairmaps temporally over scansets ($\mathbb{S}$) and over detector pairs ($\mathbb{P}$):

$$\tilde{\boldsymbol{T}}' = \sum_{\mathbb{P},\mathbb{S}} \Lambda \boldsymbol{w}^+ \boldsymbol{G}\boldsymbol{F}\boldsymbol{A}[\boldsymbol{T}]. \tag{48}$$

The matrix performing this transformation is defined as $\boldsymbol{R}'_{TT}$, where the prime indicates the apodization comes from the inverse variance of the pair sum timestream, $\boldsymbol{w}^+$. The final apodization is applied in Section 4.13.

The transformation from pair difference pairmaps to Q, U maps depends on the detector orientations during the observations. This transformation relies on an inversion of a 2×2 detector orientation matrix. We will now derive the matrix that performs this transformation.

Ignoring filtering, the pair difference timestream is found using the timestream forming matrix, A_{tj}:

$$d_t = \frac{1}{2}(\tau_t^A - \tau_t^B) = \alpha_t^- A_{tj} Q_j + \beta_t^- A_{tj} U_j. \tag{49}$$

Forming linear combinations of the pair difference timestream,

$$\begin{bmatrix} \alpha_t^- d_t \\ \beta_t^- d_t \end{bmatrix} = \begin{bmatrix} \alpha_t^- \alpha_t^- A_{tj} & \alpha_t^- \beta_t^- A_{tj} \\ \alpha_t^- \beta_t^- A_{tj} & \beta_t^- \beta_t^- A_{tj} \end{bmatrix} \begin{bmatrix} Q_j \\ U_j \end{bmatrix}, \tag{50}$$

and applying the pointing matrix, Λ, the vectors $\boldsymbol{\alpha}^-\boldsymbol{d}$ and $\boldsymbol{\beta}^-\boldsymbol{d}$ are binned into map pixels, i. At this point we coadd over scansets and detector pairs, and apply a weighting, $\boldsymbol{w}^-$, equal to

the inverse of the variance of the timestreams during a scanset:

$$\sum_{\mathbb{P},\mathbb{S}} \begin{bmatrix} \Lambda_{it} w_t^- \alpha_t^- d_t \\ \Lambda_{it} w_t^- \beta_t^- d_t \end{bmatrix} = \left(\sum_{\mathbb{P},\mathbb{S}} \begin{bmatrix} \Lambda_{it} w_t^- \alpha_t^- \alpha_t^- A_{tj} & \Lambda_{it} w_t^- \alpha_t^- \beta_t^- A_{tj} \\ \Lambda_{it} w_t^- \alpha_t^- \beta_t^- A_{tj} & \Lambda_{it} w_t^- \beta_t^- \beta_t^- A_{tj} \end{bmatrix} \right) \times \begin{bmatrix} Q_j \\ U_j \end{bmatrix}. \tag{51}$$

We invert the matrix on the right hand side of Equation (51) to compute a matrix that generates Q and U maps:

$$\begin{bmatrix} e_i & f_i \\ f_i & g_i \end{bmatrix} \equiv \left(\sum_{\mathbb{P},\mathbb{S}} \begin{bmatrix} \Lambda_{it} w_t^- \alpha_t^- \alpha_t^- \Lambda_{ti} & \Lambda_{it} w_t^- \alpha_t^- \beta_t^- \Lambda_{ti} \\ \Lambda_{it} w_t^- \alpha_t^- \beta_t^- \Lambda_{ti} & \Lambda_{it} w_t^- \beta_t^- \beta_t^- \Lambda_{ti} \end{bmatrix} \right)^{-1}, \tag{52}$$

where the t index has been summed over to find each of the elements in the 2×2 matrix on the right hand side and A_{tj} has been replaced by Λ_{ti} so the equation now determines the Q, U values in the observed map, Q_i, U_i. There is one 2×2 matrix inversion performed for each pixel, i, in the observed map. In other words, one value of e_i, f_i, and g_i is computed for each pixel in the observed map, and filled into the i-th diagonal element of $\boldsymbol{e}$, $\boldsymbol{f}$ and $\boldsymbol{g}$.

The pairmaps $\tilde{\boldsymbol{m}}_{\alpha^-}$ and $\tilde{\boldsymbol{m}}_{\beta^-}$ are transformed into Stokes Q, U by multiplying by $\begin{bmatrix} \boldsymbol{e} & \boldsymbol{f} \\ \boldsymbol{f} & \boldsymbol{g} \end{bmatrix}$. Observed Q, U maps are found according to:

$$\begin{bmatrix} \tilde{\boldsymbol{Q}}' \\ \tilde{\boldsymbol{U}}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{e} & \boldsymbol{f} \\ \boldsymbol{f} & \boldsymbol{g} \end{bmatrix} \sum_{\mathbb{P},\mathbb{S}} \mathcal{P}_{\mathbb{P},\mathbb{S}} \begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{U} \end{bmatrix}, \tag{53}$$

where $\mathcal{P}$ was defined in Equation (36).

If the sum of $\mathcal{P}$ matrices could be inverted, it would be possible to use this inverse to recover an unbiased estimate of the original Q and U. However, $\mathcal{P}$ is singular because it includes polynomial filtering and scan-synchronous signal subtraction, which completely remove some modes that were present in the original maps. We therefore use instead the matrix defined by the matrix inversion in Equation (52), which does not include these filtering operations. Even the inversion in Equation (52) is singular unless the coadded data contains observations at multiple detector angles, Ψ_t. Observations at multiple detector orientations are made through deck rotations or by coadding over receivers in different orientations. As described in Section 4.2, deck rotations occur between phases, so coadding over phases makes the matrix invertible.

Including deprojection, Equation (53) becomes:

$$\begin{bmatrix} \tilde{\boldsymbol{Q}}' \\ \tilde{\boldsymbol{U}}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{e} & \boldsymbol{f} \\ \boldsymbol{f} & \boldsymbol{g} \end{bmatrix} \sum_{\mathbb{P}} \left(\boldsymbol{D} \sum_{\mathbb{S}}^{\mathbb{S} \in \text{phase}} \mathcal{P}_{\mathbb{S}} \right)_{\mathbb{P}} \begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{U} \end{bmatrix}. \tag{54}$$

This represents the entire Q, U map making process for signal simulations: from input maps to observed maps, including filtering operations. It can be summarized as:

$$\begin{bmatrix} \tilde{\boldsymbol{Q}}' \\ \tilde{\boldsymbol{U}}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}'_{QQ} & \boldsymbol{R}'_{QU} \\ \boldsymbol{R}'_{UQ} & \boldsymbol{R}'_{UU} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{U} \end{bmatrix}. \tag{55}$$

4.12. Non-Apodized Observation Matrix

As constructed, the matrix, $\boldsymbol{R}'$, contains an apodization based on the inverse variance of the timestreams, $\boldsymbol{w}^+$ and $\boldsymbol{w}^-$. We can,

however, choose to remove this apodization, producing maps with equal weight across the field in units of μK. We construct the quantities:

$$W^+ = \sum_{\mathbb{P},\mathbb{S}} (\Lambda w^+ \Lambda^\top)_{\mathbb{P},\mathbb{S}}, \tag{56}$$

$$W^- = \begin{bmatrix} e & f \\ f & g \end{bmatrix} \sum_{\mathbb{P},\mathbb{S}} \left(\begin{bmatrix} \Lambda \alpha^- \alpha^- w^- \Lambda^\top \\ \Lambda \beta^- \beta^- w^- \Lambda^\top \end{bmatrix} \right)_{\mathbb{P},\mathbb{S}}, \tag{57}$$

and use these to remove the apodization from the observation matrix, solving for the non-apodized observation matrix, $\mathcal{R}$:

$$\mathcal{R}_{TT} = (W^+)^{-1} R'_{TT} \tag{58}$$

$$\begin{bmatrix} \mathcal{R}_{QQ} & \mathcal{R}_{QU} \\ \mathcal{R}_{UQ} & \mathcal{R}_{UU} \end{bmatrix} = (W^-)^{-1} \begin{bmatrix} R'_{QQ} & R'_{QU} \\ R'_{UQ} & R'_{UU} \end{bmatrix} \tag{59}$$

4.13. Selecting the Observation Matrix's Apodization

Using the non-apodized observation matrix of Section 4.12, we can create an observation matix with an arbitrary apodization, Z. The matrix R is constructed as follows:

$$R_{TT} = Z \mathcal{R}_{TT} \tag{60}$$

$$\begin{bmatrix} R_{QQ} & R_{QU} \\ R_{UQ} & R_{UU} \end{bmatrix} = \begin{bmatrix} Z & 0 \\ 0 & Z \end{bmatrix} \begin{bmatrix} \mathcal{R}_{QQ} & \mathcal{R}_{QU} \\ \mathcal{R}_{UQ} & \mathcal{R}_{UU} \end{bmatrix} \tag{61}$$

A sensible choice for the apodization, Z, may be the inverse variance mask we removed in Section 4.12, or a smoothed version thereof. However, there is freedom to choose any apodization at this point, and this may prove useful in joint analyses with other experiments, where the analysis combines maps with low noise regions in slightly different regions of the sky.

4.14. Summary

We have constructed a matrix, R, which performs the linear operations of polynomial filtering, scan-synchronous signal subtraction, deprojection, weighting and pointing. R has dimensions $(3\tilde{n}_p, 3n_p)$ where $\tilde{n}_p$ is the number of pixels in the BICEP map and n_p is the number of pixels in the input HEALPix map.

Using the observation matrix, the entire process of generating a signal simulation from an input map is:

$$\begin{bmatrix} \tilde{T} \\ \tilde{Q} \\ \tilde{U} \end{bmatrix} = \begin{bmatrix} R_{TT} & 0 & 0 \\ \hline 0 & R_{QQ} & R_{QU} \\ 0 & R_{UQ} & R_{UU} \end{bmatrix} \begin{bmatrix} T \\ Q \\ U \end{bmatrix}. \tag{62}$$

Here the off diagonal terms, R_{QU} and R_{UQ}, exist because the filtering operations are performed on pair difference time-streams, which are a combination of Q and U.

The deprojection operator contains regression against a temperature map that must be chosen *before* constructing the observation matrix. The deprojection operator is a linear filtering operation, and it only removes beam systematics arising from one particular temperature field. One could in principle apply the deprojection operator to Q, U maps corresponding to a different temperature field. The operator would remove the same modes from the polarization field, but these modes would not correspond to those which had been mixed between T and Q, U by beam systematics (or TE correlation).
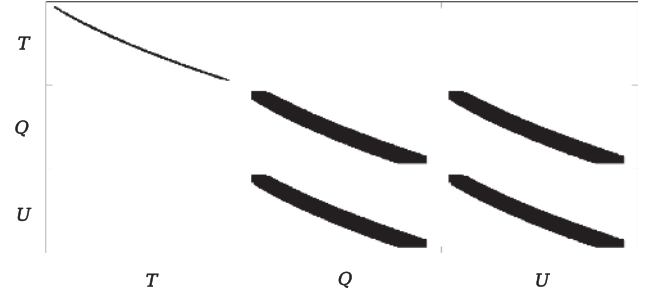

Figure 9. Observation matrix, R: filled elements of the observation matrix for the BICEP2 3 year data set. *TQ*, *TU*, *UT*, and *QT* are empty because no $T \rightarrow P$ leakage is simulated. The horizontal axis corresponds to the HEALPix pixelization and has $3 \times 111{,}593$ elements. The vertical axis corresponds to the BICEP pixelization, and has $3 \times 23{,}600$ elements. The matrix has only $\sim$5% of its elements filled.

When constructing an ensemble of E-mode realizations for use in Monte Carlo power spectrum analysis, the *TE* correlation and the fixed temperature sky force us to build constrained realizations. The ensemble of simulations all contain identical temperature fields, so we cannot use them for analysis of temperature, which is acceptable because the focus of our analysis is polarization. The ensemble does contain different realizations of Q, U, constrained for the given temperature field, and these can be used in Monte Carlo analysis of polarization. The details of constructing these constrained input maps is the subject of the Appendix.

Because the construction of R_{QQ}, R_{QU}, and R_{UU} depends on a fixed temperature field, the deprojection templates can be thought of as numerical constants. R_{TT} performs a separate filtering on the temperature field that is largely decoupled from the filtering of Q and U. To include systematics that leak temperature to polarization, the terms R_{TQ} and R_{TU} would in principle need to be non zero. However, so long as the leakage corresponded to modes being removed by the deprojection matrix D the the deprojection elements in the R_{QQ}, R_{QU}, and R_{UU} blocks would ensure that the output Q, U maps were identical.

Although the nominal dimensions of R are large, our constant elevation scan strategy means that R is only filled for pixels at roughly the same declination. This means that R is largely sparse, as shown in Figure 9.

Some intuition about the operations the observation matrix performs can be gained by plotting a column of the matrix reshaped as maps—see Figure 10. The column chosen in this case corresponds to a central pixel in the observed field. It shows how Q and U values in the observed map are sourced from a Q pixel in the HEALPix map. The bright pixel in the Q observed map corresponds to the location of the input Q. The effects of polynomial and scan-synchronous signal subtraction are visible to the left and right of the bright Q pixel. These two types of filtering are performed on scansets and are therefore confined to a row of pixels. Deprojection operates on phases, creating the effects seen at other declinations. Because all of these filtering operations are performed on pair difference data, which contains linear combinations of Q and U, signal in the observed U map can be created by signal in the input HEALPix Q map. This is why the U map in Figure 10 is non-zero.
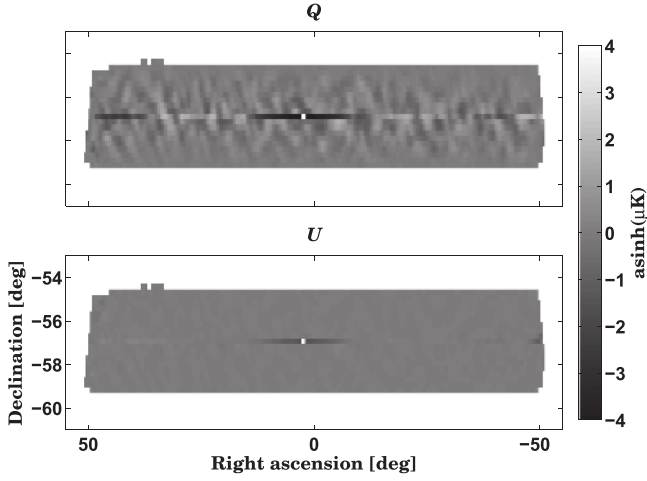
Figure 10. A single column of the observation matrix $\boldsymbol{R}$, for a `HEALPix` Q pixel near the center of our field. The value of a single input Q pixel affects both Q and U values in the observed map over the range of declinations covered in a phase.

4.15. Forming Maps from Real Timestreams

We can form observed maps from the real timestreams using the matrices constructed above:

$$\tilde{T}_{\text{real}} = \boldsymbol{Z}(\boldsymbol{W}^+)^{-1}\sum_{\mathbb{P},\mathbb{S}}(\boldsymbol{\Lambda}\boldsymbol{w}^+\boldsymbol{GFs})_{\mathbb{P},\mathbb{S}} \qquad (63)$$

$$\begin{bmatrix} \tilde{Q}_{\text{real}} \\ \tilde{U}_{\text{real}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z} & 0 \\ 0 & \boldsymbol{Z} \end{bmatrix}(\boldsymbol{W}^-)^{-1}\begin{bmatrix} \boldsymbol{e} & \boldsymbol{f} \\ \boldsymbol{f} & \boldsymbol{g} \end{bmatrix}$$
$$\times \sum_{\mathbb{P}}\left(\boldsymbol{D}\overset{\mathbb{S}\in\text{phase}}{\sum_{\mathbb{S}}}\left(\begin{bmatrix} \boldsymbol{\Lambda} & 0 \\ 0 & \boldsymbol{\Lambda} \end{bmatrix}\begin{bmatrix} \boldsymbol{\alpha}^- \\ \boldsymbol{\beta}^- \end{bmatrix}\boldsymbol{w}^-\boldsymbol{GFd}\right)_{\mathbb{S}}\right)_{\mathbb{P}}. \quad (64)$$

It is important to note that the exact same matrices are used to process the real data in Equation (64) as are used to construct the simulated maps in Equation (35).

4.16. Equivalence of Observation Matrix with Standard Pipeline

The matrix formalism described above is self-contained and complete in the sense that it contains the tools necessary to create real data maps and simulated maps.

We demand that the map making and filtering operations be identical between the standard pipeline and the observation matrix. It is straightforward to test this equivalence: simulated maps run through the standard pipeline must be identical to the maps found with the observation matrix. Figures 11 and 12 show that the two match quite well, within a few percent over the multipoles of $50 < l < 350$. The lack of a perfect match is due to the difference in deprojection timescale and because the standard pipeline uses $N_{\text{side}} = 2048$ `HEALPix` input maps that are Taylor interpolated, whereas the observation matrix uses $N_{\text{side}} = 512$ input `HEALPix` maps with nearest neighbor interpolation.

5. SIGNAL COVARIANCE MATRIX, C

The signal covariance matrix contains the pixel-pixel covariances of a map for a given spectrum of Gaussian fluctuations. The diagonal entries contain the variance of each pixel, and each row describes the covariance of a given pixel
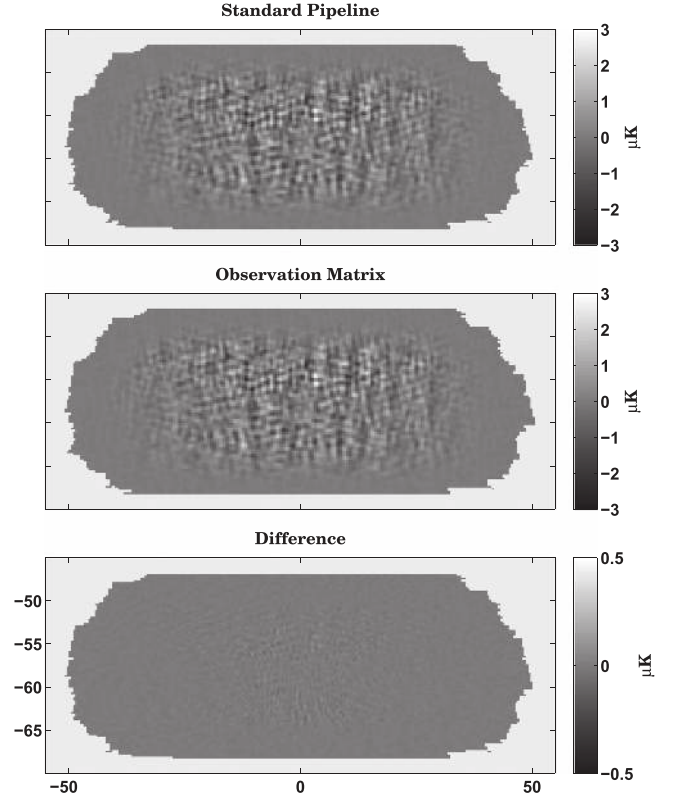

Figure 11. Comparison of observed Q maps created by the observation matrix and the standard pipeline. The input map for both is from the same simulation realization. There are small differences due the difference in deprojection timescales, input `HEALPix` map resolution, and interpolation.
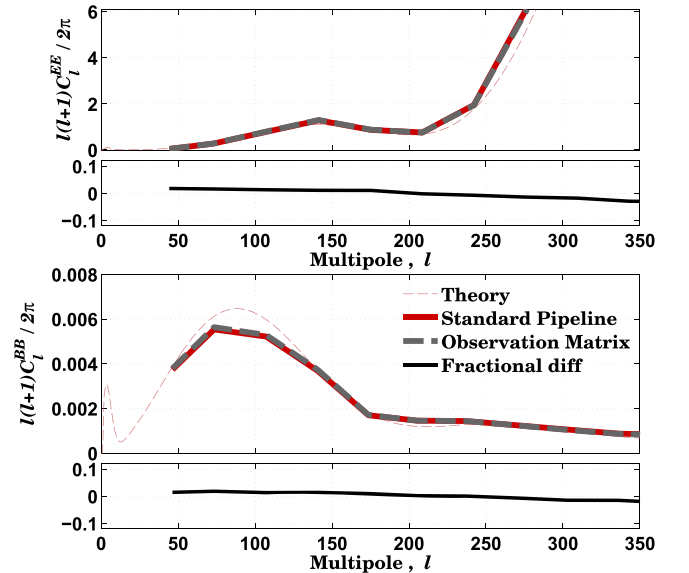

Figure 12. Comparison of power spectra of maps created by the observation matrix and the standard pipeline. The input map for both is from the same simulation realization, which differs from the theory curve for this particular realization in the Bicep field. The two methods are fractionally the same to within a few percent over the multipoles of interest, $50 < l < 350$.

with the other pixels in the map. For $[T, Q, U]$ maps, the covariance matrix contains nine sub-matrices for the correlations between T,Q, and U.

5.1. True Sky Signal Covariance Matrix

A pixel on the sky at location i, has values of the Stokes parameters:

$$x_i \equiv \begin{pmatrix} T_i \\ Q_i \\ U_i \end{pmatrix}. \qquad (65)$$

The 3×3 pixel-pixel covariance between two locations on the sky, i and j, is given by:

$$\boldsymbol{C}_{i,j} \equiv \langle x_i x_j \rangle = \mathbb{R}(\alpha) \boldsymbol{M} (r_i \cdot r_j) \mathbb{R}(\alpha)^\top. \qquad (66)$$

The covariance matrix, $\boldsymbol{M}$, is defined with the Q, U convention referenced to the great circle connecting the two points, i, j. For a particular spectrum $\boldsymbol{M}$ depends only on the dot product between the pixels, $r_i \cdot r_j$. $\boldsymbol{M}$ contains nine symmetric sub-matrices:

$$\boldsymbol{M} (r_i \cdot r_j) = \begin{pmatrix} \langle T_i T_j \rangle & \langle T_i Q_j \rangle & \langle T_i U_j \rangle \\ \langle Q_i T_j \rangle & \langle Q_i Q_j \rangle & \langle Q_i U_j \rangle \\ \langle U_i T_j \rangle & \langle U_i Q_j \rangle & \langle U_i U_j \rangle \end{pmatrix}. \qquad (67)$$

The 3×3 matrix, $\mathbb{R}$, is applied to rotate from this local reference frame to a global frame where Q, U are referenced to the North–South meridians. The angle between the great circle connecting any two points and the global frame is given by the parameter α

$$\mathbb{R}(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha \\ 0 & -\sin 2\alpha & \cos 2\alpha \end{pmatrix}. \qquad (68)$$

Changing the sign of α allows us to change the polarization convention from IAU to HEALPix (U to $-U$), see Hamaker & Bregman (1996). We have chosen to use the IAU convention for BICEP2 and Keck Array covariance matrices.

The true sky pixel-pixel signal covariance matrix for the Stokes Q, U parameters is derived in Kamionkowski et al. (1997), Zaldarriaga (1998). To calculate the covariances, we follow some of the suggestions in Appendix of Tegmark & de Oliveira-Costa (2001).

We use the HEALPix ring pixelization, which allows the covariance to be calculated simultaneously for all pixels at a particular latitude that are separated by the same distance.[23] This shortcut is exploited by simultaneously calculating all equidistant pixels for rows in the map that have the same latitude to within 1×10^{-4} degrees. This approximation is much smaller than the $\sim$7 arcminute pixels in $N_{\rm side} = 512$ maps, and the rounding error has been found to be insignificant.

5.2. Observed Signal Covariance Matrix, $\tilde{C}$

The observed signal covariance matrix contains the pixel-pixel covariance in the observed BICEP pixelized maps. Theoretically, modifying the true sky signal covariance matrix is simple: take the unobserved signal covariance of $\boldsymbol{C}$ of Section 5.1 and the observation matrix $\boldsymbol{R}$ from Section 4 and form the product:

$$\tilde{C} = \boldsymbol{R} \boldsymbol{C} \boldsymbol{R}^\top. \qquad (69)$$

---

[23] This quality of the HEALPix maps is by design, see Gorski et al. (2005).
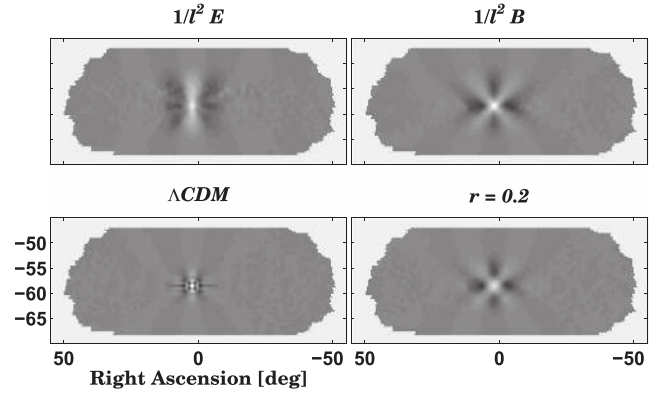

Figure 13. Maps showing a row of the observed covariance matrix $\tilde{C}$. The row selected corresponds to the covariance of an individual Q pixel at the center of the map. The top row shows the covariance used to calculate the pure E and B-modes described in Section 6. The bottom row shows the covariance for an input spectrum corresponding to ΛCDM (left), and $r = 0.2$ tensors (right).

This equation results in a symmetric positive definite matrix, which is rank deficient because of the filtering steps in the observing process.

Unfortunately, performing this multiplication is computationally demanding: The input $\boldsymbol{C}$ is a square matrix, with $3 \times 111{,}593$ elements on a side, corresponding to the elements of T, Q, and U. To reduce the memory requirements of the calculation, we divide the covariance matrix, $\boldsymbol{C}$, into row subsets and calculate in parallel. Once a row subset is calculated, the observation matrix is immediately applied to transform the HEALPix covariance to the observed map covariance, which reduces the dimensions of the covariance to the 23,600 pixels of the observed maps.

The covariance and observed covariance should both be symmetric, which provides a good check on our math. Usually the output is slightly (fractionally, $\sim 1/10^7$) non-symmetric due to rounding errors in the multiplication, and we force the final matrix to be symmetric to numerical precision by averaging across the diagonal before moving to the next steps, since symmetric matrices often allow the use of faster algorithms.

A row of the observed covariance matrix can be reshaped into a map, which reveals the structure of the covariance for a particular pixel, see Figure 13.

6. E/B SEPARATION USING A PURIFICATION MATRIX

The observed covariance matrix contains expected pixel-pixel covariance in our observed maps given an initial spectrum. The observation matrix, $\boldsymbol{R}$, has made the E and B-mode spaces of the observed covariance non-orthogonal. The result of Section 3 is that we can find the orthogonal pure E and pure B spaces by solving the eigenvalue problem from Equation (17): $\tilde{C}_B x_i = \lambda_i \tilde{C}_E x_i$.

6.1. Construction of Purification Matrix

As written, Equation (17) is not solvable: $\tilde{C}_B$ has a null space that is the set of pure E-modes. Similarly, the space of pure B-modes is the null space of $\tilde{C}_E$. By adding the identity matrix multiplied by a constant, $\sigma^2 \boldsymbol{I}$, to the covariance matrices we regularize the problem to find approximate solutions and eliminate the null spaces:

$$(\tilde{C}_B + \sigma^2 \boldsymbol{I}) x_i = \lambda_i (\tilde{C}_E + \sigma^2 \boldsymbol{I}) x_i. \qquad (70)$$
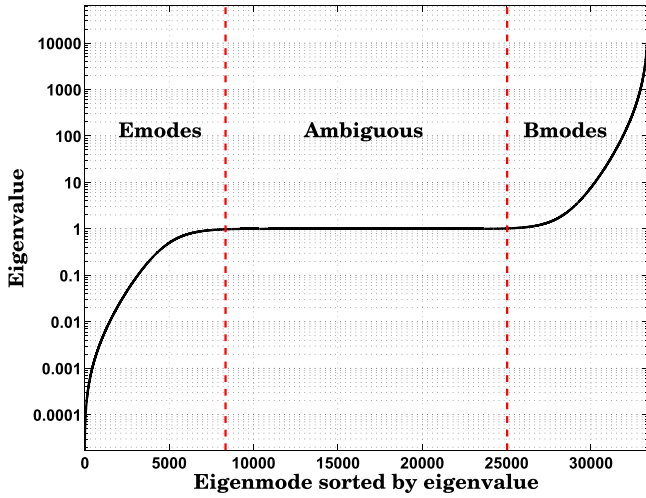
Figure 14. The generalized eigenvalues for the BICEP2 observed covariance matrix, sorted by magnitude. Eigenvalues near one correspond to ambiguous modes: the modes that are simultaneously E and B in the observed space and must be thrown out. By selecting eigenmodes with eigenvalues that are the largest and smallest $1/4$ of the set of eigenvalues (shown to the left and right of the dashed red lines), we can construct subspaces that span the spaces of B-modes and E-modes that can be effectively observed using our scan strategy and analysis.

The amplitude of σ^2 sets the relative magnitude of the ambiguous mode eigenvalues versus the pure E and pure B-mode eigenvalues. The eigenvalues are shown in Figure 14. In our analysis, we choose σ^2 to be $1/100$ the mean of the diagonal elements of the covariance matrices, $\tilde{C}_E$ and $\tilde{C}_B$.

In Equation (70), $\tilde{C}_E$ is an observed covariance matrix for $[Q, U]$ that is constructed according to Equations (11) and (69). The input spectrum is set to a steeply red E-mode spectrum, $C_l^{EE} = 1/l^2$, $C_l^{BB} = 0$. $\tilde{C}_B$ is the same except for an input spectrum with only B-modes, $C_l^{BB} = 1/l^2$, $C_l^{EE} = 0$. The eigenmodes in x_i with the largest eigenvalues comprise a set of vectors that span a space of pure B-modes, b_i. The pure B quality of these vectors can be seen in the fact that the product $\tilde{C}_E x$ is much smaller than the product $\tilde{C}_B x$. The eigenmodes in x_i with the smallest eigenvalues comprise a set of vectors that span a space of pure E-modes, e_i.

Using the reddened input spectrum $1/l^2$ causes the magnitude of the eigenvalues to be proportional to the band of multipole, l, that each mode contains. The steepness of the spectrum ensures each mode contains power from a limited range of l. The particular choice of $1/l^2$ is arbitrary.

A basis constructed from a subset of eigenmodes with large eigenvalues spans a subspace of pure B-modes. We arbitrarily choose the pure B subspace to consist of eigenmodes whose corresponding eigenvalues are the largest $1/4$ of the set. We find modes contained in this subset adequate for preserving power up to $l \sim 700$. The pure E subspace is similarly constructed with eigenmodes corresponding to the $1/4$ smallest eigenvalues. Figure 14 shows the sorted eigenvalues, and Figure 15 shows four eigenmodes of the BICEP2 observed covariance.

Using the set of pure E and pure B-mode basis vectors, two projection matrices are constructed from the outer products:

$$\Pi_E = \sum_i e_i e_i^\top$$
$$\Pi_B = \sum_i b_i b_i^\top, \tag{71}$$

which we call the purification matrices for pure E and pure B. Operating the purification matrices on an input map projects onto the space of pure E and pure B:

$$\tilde{m}_{\text{pureE}} = \Pi_E \tilde{m}$$
$$\tilde{m}_{\text{pureB}} = \Pi_B \tilde{m}. \tag{72}$$

From the construction of the pure B basis, b_i, one can see that the vector $\tilde{m}_{\text{pureB}}$ vanishes for arbitrary input containing only E-modes: $\tilde{m} = R \cdot \sum a_{lm}^E Y_{lm}^E$, as desired for a purified B map.

7. MATRIX E/B SEPARATION APPLIED TO BICEP2

This section describes the application of the matrix based E/B separation to the BICEP2 data set. The technique relies on the existence of the observation matrix and purification matrix from the previous sections.

7.1. Motivation for Matrix Based E/B Separation in BICEP2

The BICEP2 and Keck Array analysis pipeline contains the following attributes that can leak E-modes to B-modes: partial sky coverage, timestream filtering (including deprojection), and choice of map projection+estimator. A simulation demonstrating the leaked B-mode maps for each of these effects is shown in Figure 16. These maps are created by applying the standard E and B estimator in Fourier space and then an inversion back to map space. The observations and data reduction produce three classes of E/B leakage:

1. *Apodization*: The first obvious deviation from the ideal full sky map is the partial sky coverage of BICEP2 and Keck Array maps. Once a boundary is imposed E/B leakage is created. Map boundary effects are reduced by an apodization window which tapers the maps smoothly to zero near the edges. Use of apodization windows is common practice in any Fourier transform analysis of finite regions to prevent "ringing" near boundaries. For small regions of sky, effects of the map boundary dominate the leakage even after apodization is applied. The apodization window used in BICEP2 and the Keck Array are modified inverse variance maps. We apply a smoothing Gaussian with a width of $\sigma = 0°.5$ to the inverse variance map, and, for the combined analyses of BICEP2 and the Keck Array, we use a geometric mean of the individual experiments' apodization maps.

2. *Map projection*: The total extent of the BICEP2 and Keck Array maps is about $50°$ on the sky in the direction of R. A., over a declination of $-70 < \delta < 45$. The chosen projection is a simple rectangular grid of pixels in R.A. and decl.. Taking standard discrete Fourier transforms of such maps results in significant E/B leakage. While we note that other map projections will have significantly lower distortion, such effects will be present for all projections when subjected to Fourier transform. However, note that the A and Λ matrices together fully encode the mapping from the underlying curved sky to the flat sky of the observed map pixels, and therefore so does the observing matrix R derived from these.

3. *Linear filtering effects*: There are three main analytic filters applied in the standard pipeline: polynomial filter, scan-synchronous subtraction, and deprojection. With respect to E to B leakage, all three filters are similar: by
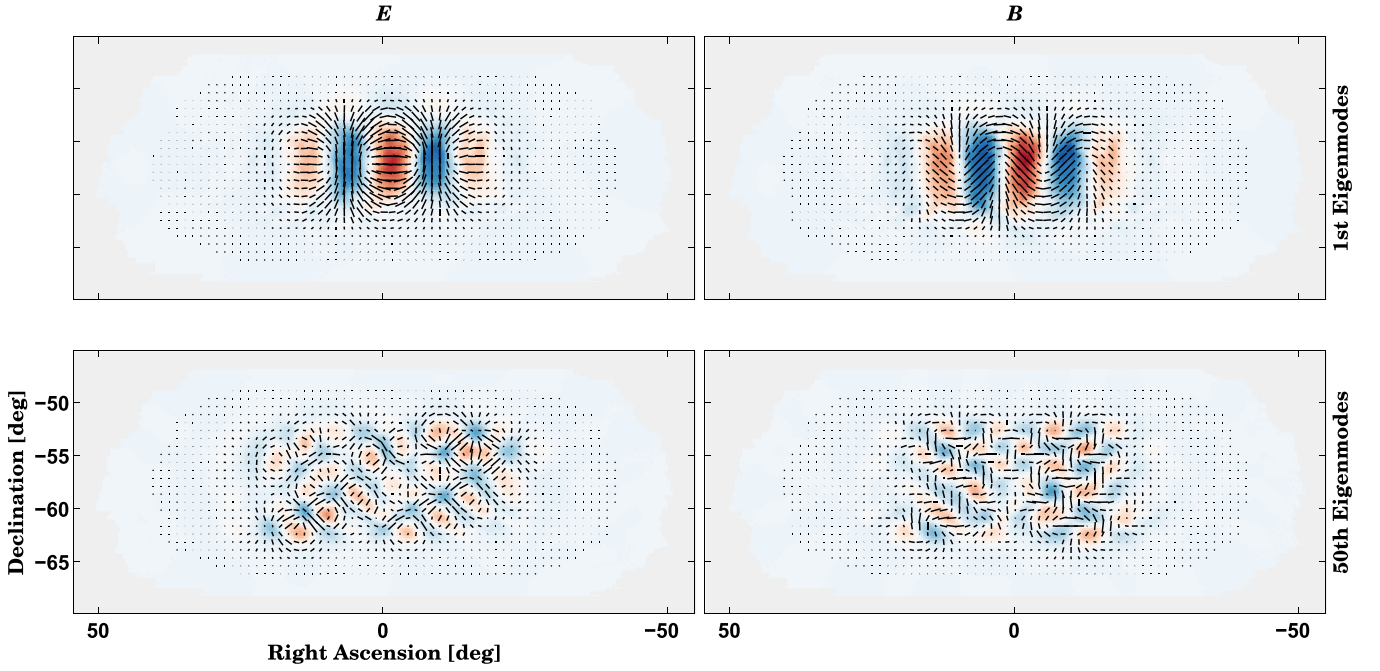
Figure 15. Eigenmodes of the BICEP2 observed covariance matrix. Shown are the modes corresponding to the largest and 50th largest eigenvalues of Equation (70). Colormap shows amplitude of *E* and *B*-modes. The eigenvalues are shown graphically in Figure 14.
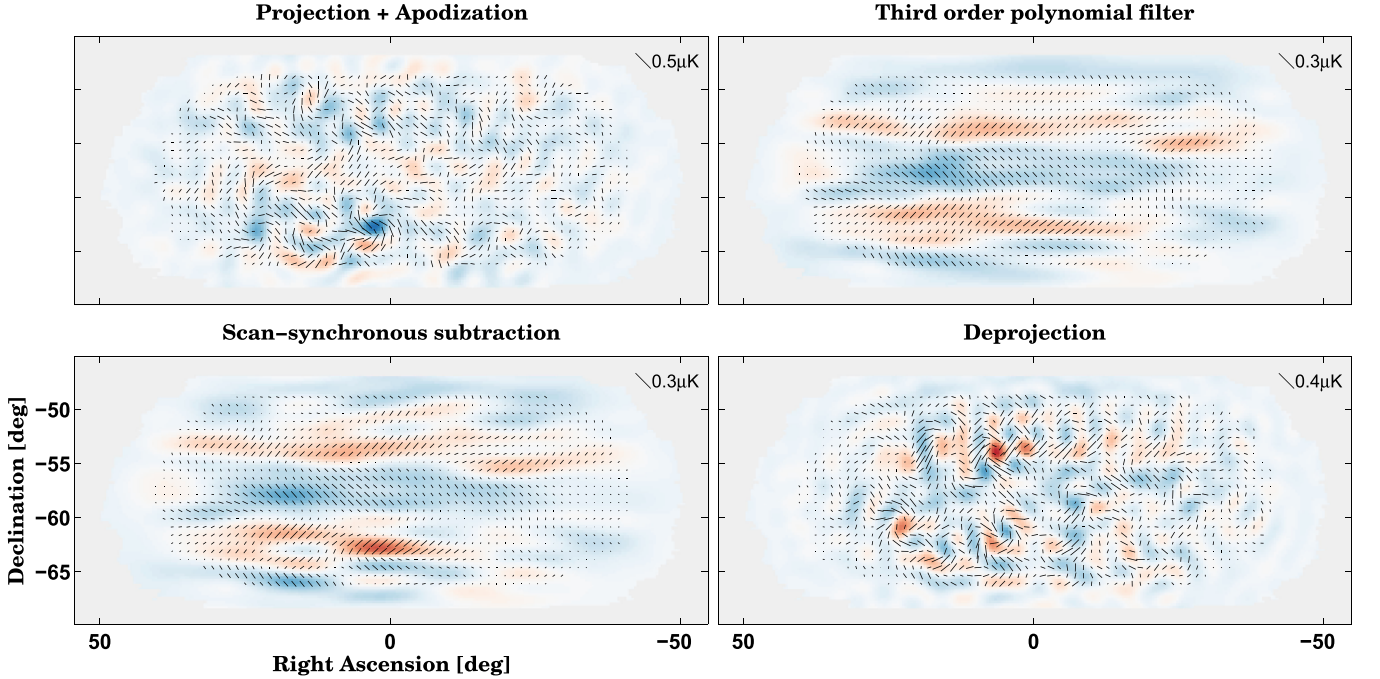

Figure 16. *E* to *B* leakage maps: examples of leaked *B*-modes in the BICEP2 maps. Top row, left: leaked *B*-modes due to map projection and apodization. Top row, right: leaked *B*-modes due to third order polynomial subtraction. Bottom row, left: leaked *B*-modes due to scan-synchronous signal subtraction. Bottom row, right: leaked *B*-modes due to deprojection of beam systematics.

removing modes in the *Q*, *U* maps, an *E*-mode can be turned into an observed *B*-mode. Polynomial filtering and scan-synchronous signal subtraction create comparable leakage, both in amplitude and morphology. Deprojection creates more power in the leaked *B* maps, and at smaller angular scales, than either polynomial filtering or scan-synchronous signal subtraction.

If matrix purification is not used, the sample variance of *E*/*B* leakage in the BICEP2 *BB* power spectrum is comparable to the uncertainty due to instrumental noise. This is clearly highly undesirable, and led to the development of the purification matrix described in this paper. The purification matrix "knows about" all of the *E*/*B* mixing effects and how to deal with them.

7.2. Effectiveness of Purification Matrix

The effectiveness of the purification matrix given by Equation (72) can be immediately tested by applying the operator to a vector of [*Q*, *U*] maps simulated with the standard
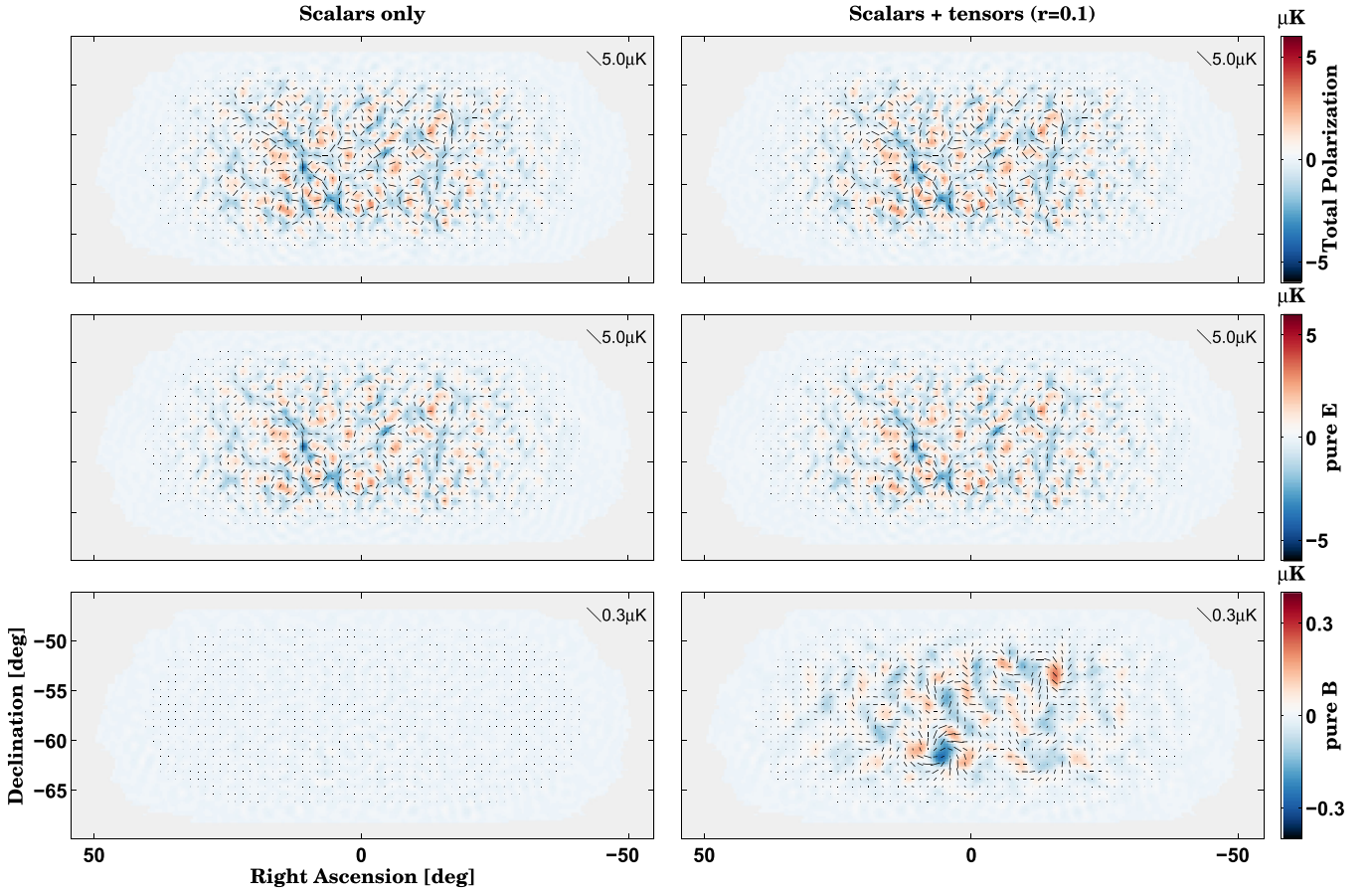
14

Figure 17. Polarization maps showing the effectiveness of the BICEP2 purification matrix at separating noiseless simulations into pure E and pure B. Left column: on a scalar only unlensed-ΛCDM BICEP2 simulation. Right Column: on the same simulation with the addition of a small tensor component. Top Row: the total polarization of the BICEP2 observed map, containing E and ambiguous modes. Center Row: pure E-mode map, constructed by projecting the total polarization map onto the E eigenmodes found in Equation (70). Bottom Row: pure B-mode map, constructed by projecting the total polarization map onto the B eigenmodes of Equation (70).

pipeline. The upper left map of Figure 17 shows an observed map whose input is unlensed-ΛCDM E-modes, and the next two rows in that column show the resulting E and B maps after projection onto the pure E and B spaces. The right column of Figure 17 shows an observed input map with both unlensed-ΛCDM and $r = 0.1$ projected onto pure E and B-modes.

Matrix purification is integrated with the existing BICEP analysis code by applying the purification operator to maps before calculating the power spectra. Since the observation matrix is only used for this purification step, the purification matrix need only work well enough to result in E/B leakage less than the noise level of the experiment. Therefore, it is acceptable to use an approximate observation matrix constructed from a subset of the full observation list, as long as it is representative of the full scan strategy. This shortcut was employed in Keck Array & BICEP2 Collaborations V (2015), but the results shown in this section from BICEP2 are from the full set of observations.

Figure 18 compares the spectra for purified maps to the spectra from maps without purification, and to the spectra found using the improved estimator suggested in Smith (2006). Both the purified maps and unpurified maps use the standard E and B estimator in Fourier space, and we refer to the unpurified maps processed this way as the "normal method." Figure 18 shows the spectra for 200 noiseless unlensed-ΛCDM simulations passed through the three estimators. The leaked power is roughly three orders of magnitudes smaller when using the
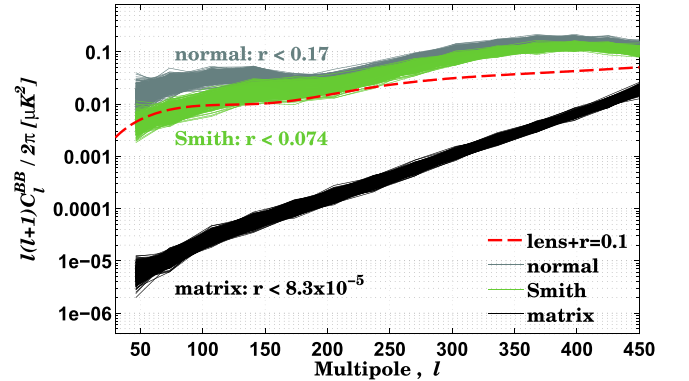

Figure 18. BB power spectra of noiseless unlensed-ΛCDM ($r = 0$) simulations, estimated using various methods, demonstrating the effectiveness of the BICEP2 purification matrix. The E/B leakage using the matrix estimator is 3 orders of magnitude lower than other methods.

matrix purification than when using the normal method or Smith estimator. While the Smith estimator improves over the normal estimator by eliminating E/B leakage from apodization, it does not account for spatial filtering, which is a significant source of E/B leakage in the analysis pipeline. The mean of the leaked power is de-biased from the final power spectra, so what matters is the variance of the leaked spectra. Computing the 95% confidence limits based on the variance in each of the
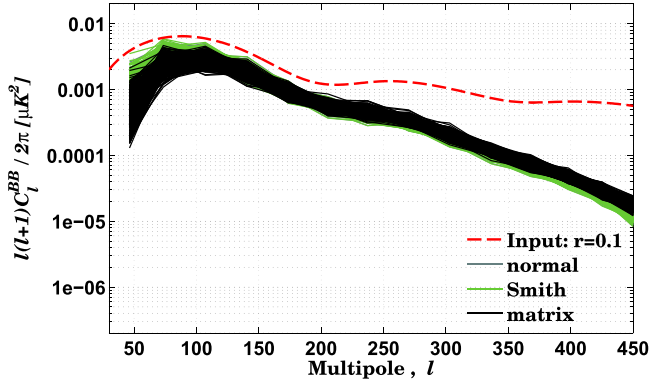
Figure 19. *BB* power spectra of noiseless unlensed ($r = 0.1$) tensor only simulations, estimated using various methods. All methods suffer from loss of power due to filtering and beam effects. The removal of ambiguous modes at low *l* results in a further decrease in power for the matrix method. Note that the spectra in this plot have not been corrected for the beam and filter suppression factors, but in Figure 18 the correction is applied.

three methods, we find that in the absence of *B*-mode signal or instrumental noise, the matrix estimator achieves a limit on the tensor-to-scalar ratio of $r < 8.3 \times 10^{-5}$, while the normal method and Smith estimator achieve limits of $r < 0.17$ and $r < 0.074$ respectively.

Figure 19 shows the spectra from the three estimators for input maps containing only input *B*-modes at the level of $r = 0.1$. The spectra for all three estimators show beam roll off at high *l*. At the lowest *l*, the filtering prevents large angular scale modes from being measured. For multipoles around $l \sim 100$, the matrix estimator recovers slightly less signal than the other two methods. However, the extra power measured by the other methods near $l \sim 100$ largely comes from the ambiguous modes. On the left of Figure 18, these ambiguous modes are seen as the bump in the normal and Smith method near $l \sim 100$.

The total number of degrees of freedom in each band power can be estimated according to the formula:

$$N_{l'} = 2 \frac{(m_{l'})^2}{\sigma_{l'}^2}, \tag{73}$$

where $m_{l'}$ is the mean of the simulations in band power l' and $\sigma_{l'}^2$ is the variance of the simulations in the band power. Table 1 shows the number of degrees of freedom for the three estimators for a tensor *B*-mode. The fewer degrees of freedom at low *l* for the matrix estimator are consistent with the decrease in recovered power on the left side of Figure 18. The highest *l* bins in Table 1 also show fewer degrees of freedom in the matrix estimator because the purification matrix includes a limited number of pure eigenmodes, as shown in Figure 14.

Figure 20 shows signal plus noise spectra for a set of 200 unlensed-ΛCDM+noise spectra. The noise simulations are the standard BICEP2 sign flip realizations discussed in BICEP2 Collaboration I (2014). In Figure 20, the mean noise level and leaked *BB* power are de-biased. The resulting ensemble of simulations is used to construct the errorbars in the final spectra. The tighter distribution of the matrix estimator simulations is a result of the decrease in *E/B* leakage. Using the matrix estimator results in an improvement in the *r* limit, for BICEP2 noise level and filtering in the absence of *B*-mode signal, of about a factor of two over the Smith method. The

Table 1
Degrees of Freedom in Binned *BB* Power Spectra for Different Estimators

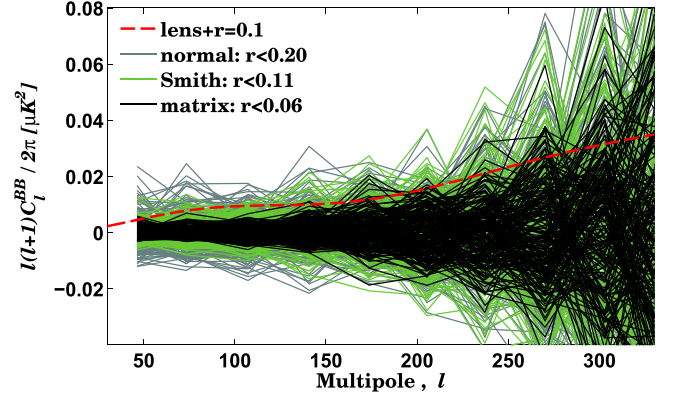| Bin Center, *l*	Degrees of Freedom		
	Normal	Smith	Matrix
37.5	12.9	15.5	8.6
72.5	40.9	41.4	34.8
107.5	71.4	69.7	66.8
142.5	83.7	81.2	81.7
177.5	120.6	116.4	116.9
212.5	156.1	153.0	141.9
247.5	172.0	172.9	145.8
282.5	202.8	200.8	177.4
317.5	189.0	185.7	155.0

Figure 20. *BB* power spectra of unlensed-ΛCDM ($r = 0$) + BICEP2 noise simulations, estimated using various methods, demonstrating the effectiveness of the BICEP2 purification matrix. For BICEP2 noise levels the constraint on *r* (in the absence of signal) is improved by about a factor of two. (The mean of the noise and leakage have been de-biased in each case.)

remaining variance in the BB spectrum of the matrix estimator is instrumental noise.

7.3. Transfer Functions

The observation matrix transforms an input HEALPix map into an observed map with a simple matrix multiplication. The speed of the operation facilitates the calculation of analysis transfer functions, which are a necessary component of the pseudo-C_l MASTER algorithm (Hivon et al. 2002).

We start with input maps, $\boldsymbol{m}_l$, which are delta functions in a particular multipole. These maps are then observed using the matrix $\boldsymbol{R}$. The spectra calculated from these maps represent the response in our analysis pipeline to the input delta function, in a manner conceptually analogous to Green's functions.

Our procedure uses two sets of HEALPix maps, one set corresponding to $TT = TE = EE = 1$ and one set with $TT = BB = 1$. The observation matrix is used to create maps for $l = 1$ through 700, with 100 random realizations for each *l*. Processing the 140,000 maps would be infeasible without the observation matrix, but using the observation matrix it can be accomplished in a few hours.

7.3.1. Band Power Window Functions

The power spectra of the output maps for a particular *l* are averaged over the $N = 100$ realizations. The averaged spectra are used to form a band power window function, $\mathcal{M}_{ll'}^{XX}$, for a particular band power, l', which is a function of the input
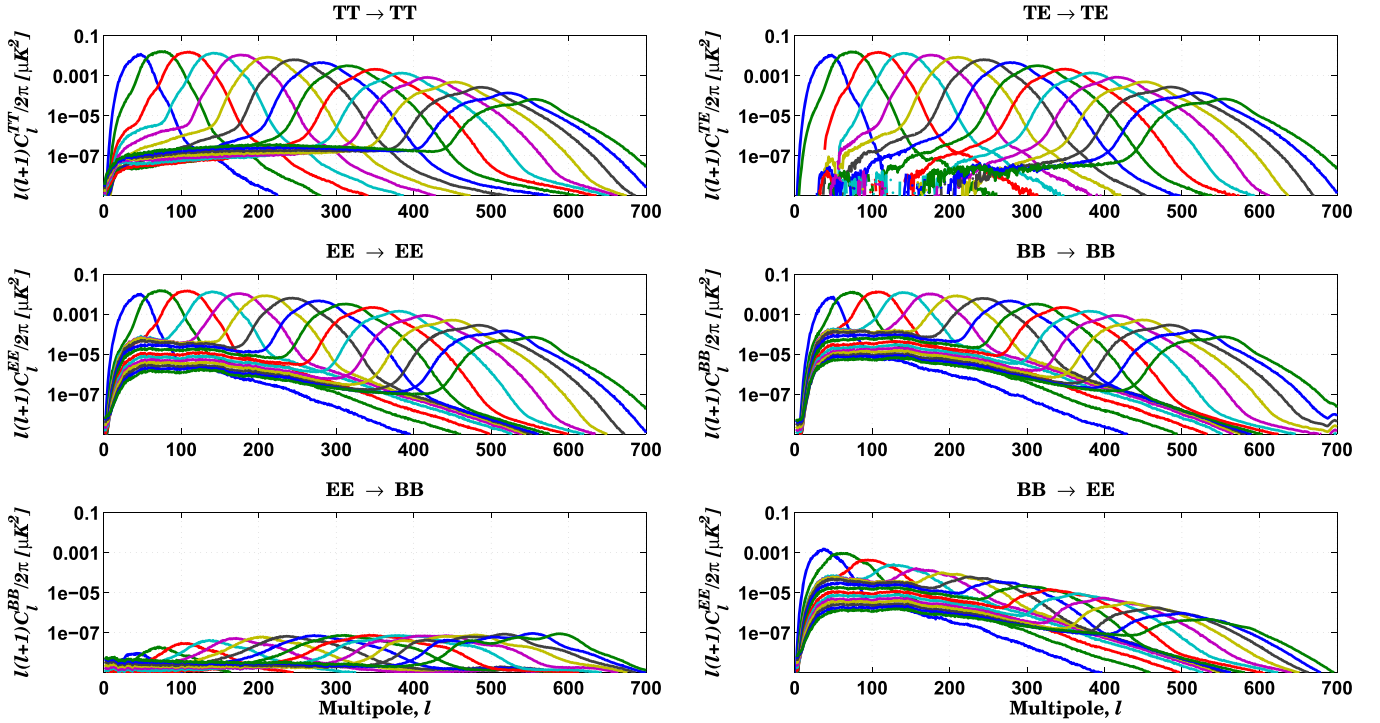
Figure 21. Band power window functions, $\mathcal{M}_{ll'}^{XX}$. Filtering causes mixing of power from low multipoles up to higher multipoles. As noted in Section 7.3.1, the $BB \rightarrow EE$ panel in the bottom right shows significantly more power since matrix purification is not applied to E-modes.

multipole of the delta function, l:

$$\mathcal{M}_{ll'}^{XX} = \frac{\sum_{r=1}^{N} \mathcal{F}_{ll'}(\boldsymbol{R}\boldsymbol{m}_l)}{N}, \qquad (74)$$

where $\mathcal{F}_{ll'}$ is the analysis pipeline's transformation from map to power spectra[24] and $XX = \{TT \rightarrow TT, TE \rightarrow TE, EE \rightarrow EE, EE \rightarrow BB, BB \rightarrow BB, BB \rightarrow EE\}$. When the input maps contain E-modes, and we measure the BB spectra, the result is the $EE \rightarrow BB$ band power window. The use of the purification matrix prevents leakage and makes these band power windows have much lower amplitude than the $EE \rightarrow EE$ or $BB \rightarrow BB$ ones. In BICEP2, it has been standard procedure not to use the E-mode purification matrix, hence the $BB \rightarrow EE$ band power window contains (irrelevant) leakage from B into E.

Calculating the band power window functions in this way accounts for all aspects of our instrument and analysis: beam convolution, sky cut, map projection, polynomial filtering, scan-synchronous signal subtraction, deprojection, and power spectrum estimation.

Figure 21 shows the results of the calculation. Although we bin in annular rings of the two-dimensional power spectra, the filtering operations move power from other l into those bins. This means the band powers are sensitive to a broader range in l than the nominal range of multipoles. The broad shelf in the band power window functions at lower l is due to filtering.

Table 2 shows the "nominal" and "measured" centers and edges of the band power bins, where "nominal" refers to the defined range of annular rings in the two-dimensional power spectra and "measured" refers to the center and $\pm1\sigma$ range

found in the end to end calculation discussed in this section. The center is the mean of the bandpower window function and the $\pm1\sigma$ interval corresponds to the percentiles of the bandpower window function between 16% and 84%.

7.3.2. Suppression Factor

The integrated area under the curve of each band power window function represents the response of that band power measurement to an input spectrum. We call this set of values the suppression factor, $S_{l'}$, since they approximate how our analysis pipeline suppresses power.

The suppression factor is plotted in Figure 22. At small angular scales, the suppression factor is dominated by the roll-off of BICEP2's 31 arcminute beam. The measured array average 150 GHz beam function is shown as the dotted line in Figure 22. The "map window function," which includes the finite size of the map and the pixel window function for the $\sim0°.25$ pixels, is shown as the dashed line.[25] At high l, the pixel window function is sub-dominant to the beam window function. At low l the suppression factor is dominated by the timestream filtering effects.

7.4. Computing Challenges

Building the observation matrix requires constructing, multiplying, and finally averaging a large number of sparse matrices. Applying the observation matrix to the true sky signal covariance matrix requires a large matrix multiplication to find the product, $\boldsymbol{R}\boldsymbol{C}\boldsymbol{R}^\top$.

Matrix multiplication is the dominant contributor to computation time. We use matrix multiplication routines built in to MATLAB. These routines incorporate a number of

---

[24] Including the steps: apply matrix purification, two-dimensional Fourier transform, construction of E and B, and binning to one-dimensional spectra.

[25] Calculated according to http://healpix.jpl.nasa.gov/html/intronode14.htm.

Table 2
BB Band Power Widths

Bin Number	Nominal			Measured		
	Low	Center	High	Low	Center	High
1	20.0	37.5	55.0	37.0	46.4	54.0
2	55.0	72.5	90.0	59.0	73.4	86.0
3	90.0	107.5	125.0	92.0	107.2	121.0
4	125.0	142.5	160.0	125.0	140.7	157.0
5	160.0	177.5	195.0	158.0	173.7	192.0
6	195.0	212.5	230.0	189.0	205.4	227.0
7	230.0	247.5	265.0	220.0	237.2	262.0
8	265.0	282.5	300.0	253.0	270.2	298.0
9	300.0	317.5	335.0	285.0	302.9	333.0

Note. Nominal and measured centers and edges of the band power bins. The measured values are extracted from the band power window functions shown in Figure 21. The low/high values for the latter are the $\pm 1\sigma$ points.
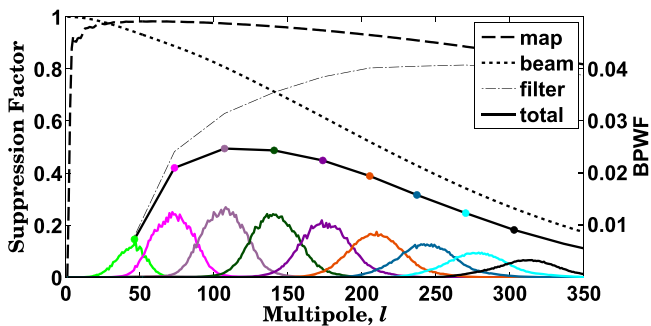

Figure 22. The *BB* suppression factor. At high *l* the beam function dominates. At low *l* the effects of filtering dominates. The map window function for the BICEP 0°.25 square pixels and finite map size is shown as a dashed line. The band power window functions are plotted in colors corresponding to individual band powers on a different scale.

optimized algorithms for computing matrix products, including BLAS (Lawson et al. 1979). At this time we have not compared the run time on GPUs with that on CPUs, but are aware of this as a possible avenue for reducing the compute time.

Solving the eigenvalue problem using the MATLAB function `eig()` takes about 48 hr and 80 GB of RAM for the full BICEP2 observed covariance. For BICEP2, this is small fraction of the total computation time. However, for experiments whose maps contain more pixels, the difficulty of the eigenvalue problem increases. In these cases, the use of distributed memory parallel code may be necessary.

For the BICEP2 results, we used computing resources provided by the Odyssey cluster at Harvard.[26] The Odyssey cluster contains 54K CPUs with 190 terabytes of RAM and 10 petabytes of storage. High memory nodes have access to 256 GB of RAM, which is useful for large matrix multi-plications. Odyssey uses SLURM[27] as its queue manager, allowing our analysis to utilize the large number of cores available.

Although the raw BICEP2 data set comprises roughly 3 TB of data, the data products from the steps in the matrix analysis chain use 17 TB of storage. Processing during the BICEP2 matrix analysis steps required roughly 1 million CPU hours. This represents a significant portion of the total computing

demand of the entire BICEP2 analysis effort, which comprised roughly 6 million CPU hours.

8. CONCLUSIONS

We have described a method for decomposing an observed polarization field into orthogonal components coming from celestial *E* and *B*-modes. The method relies on numerically calculating an observation matrix. In our case the observation matrix encodes the mathematical steps translating the true sky to an observed map including polynomial filtering, scan-synchronous subtraction, pointing of individual detectors, and linear regression of beam systematics.

Applying the observation matrix to pixel-pixel covariance matrices for *E* and *B*-modes transforms the true sky covariance into the observed space. We then solve for the *E* and *B* eigenmodes and select those modes that are orthogonal. In this way, the orthogonality relationship of the true sky is translated to the observed maps. The method accounts for all types of *E/B* leakage: boundary effects, polynomial filtering, linear regression, etc.—as long as these properties of the observing strategy and analysis have been encoded in the observation matrix, making the method more general than the method presented in Smith (2006), which only accounts for boundary effects.

The observation matrix has many other possibilities and in principle allows construction of fully optimal analyses through to power spectra or cosmological parameters for a single experiment, or a combination of experiments with partially overlapping sky coverage. One simple application which we have explored is to use the observation matrix to directly produce simulated observed maps from input maps in a single step. This is dramatically faster than the previous standard pipeline. However production of observing matrices is sufficiently costly that we have not generated them for the many alternate "jackknife" data split maps and so for the present standard simulations are still required.

We find that the matrix based *E/B* separation performs quite well, limiting the leakage to a level corresponding to $r < 1 \times 10^{-4}$, well below the noise level for any foreseeable CMB experiment. The method should prove useful for future ground based or balloon based experiments focused on measuring large angular scale *B*-modes. Experiments measuring the lensing potential using CMB polarization rely on cleanly separating *E* and *B* as well and may find the technique useful. Additionally, the ongoing search for the imprint of gravitational waves in CMB polarization will require mitigation of lensing signal from intervening structure and foreground removal, both of which are improved by cleanly separating *E* and *B*.

[26] https://rc.fas.harvard.edu/odyssey/
[27] http://slurm.schedmd.com/

research would not have been possible. We thank all those who have contributed efforts to the BICEP/Keck Array series of experiments, including the BICEP1 and BICEP3 teams.

APPENDIX
GENERATION OF CONSTRAINED REALIZATION
HEALPix MAPS

An ensemble of signal only simulations is needed for a MASTER pseudo-C_l analysis. We use CAMB with input *Planck* parameters to construct power spectra, C_l. The power spectra are used to generate high resolution HEALPix maps that serve as the starting point for each realization of the signal simulations. Lacking any constraints on the realizations, the maps generated from the power spectra will vary for both the temperature and polarization fields. However, *Planck* has measured the temperature field of the CMB to high signal to noise. Since the goal of BICEP2 and the Keck Array is to measure the polarization sky and not the temperature sky, our ensemble of simulations does not need to contain variation in the well measured temperature sky. We therefore use the temperature field of the CMB measured by *Planck* as the template for constrained realizations of the polarization field.

We have developed a technique for creating realizations of the *E*-mode sky consistent with the known *TE* correlation and the measured temperature field. These constrained realizations have been shown to contain the same *TE* correlations and *EE* spectra over many realizations of the temperature sky. However, any particular set of constrained realizations based on one temperature sky has a slightly different distribution of *TE* and *EE* than the full ensemble average.

The primary motivation for fixing the temperature field is to make the deprojection operation discussed in Section 4.10 into a linear operator. Recall that deprojection involves a regression against a fixed template of the temperature sky. If the temperature sky varies from realization to realization, deprojection becomes non-linear and cannot be expressed as a matrix operation.

A.1. TE Correlation

We start by fixing the coefficients of the spherical harmonics of the temperature sky, a_{lm}^T. The a_{lm} for the constrained *E*-modes is found in Dvorkin et al. (2008) and can be derived following the steps of a simple Cholesky decomposition. The 2×2 covariance between *T* and *E* for each mode (l, m) is:

$$\begin{bmatrix} C_l^{TT} & C_l^{TE} \\ C_l^{TE} & C_l^{EE} \end{bmatrix}. \tag{75}$$

The lower triangle Cholesky decomposition of this 2×2 matrix determines the variance and covariance in each mode:

$$\begin{bmatrix} \sqrt{C_l^{TT}} & 0 \\ C_l^{TE}/\sqrt{C_l^{TT}} & \sqrt{C_l^{EE} - (C_l^{TE})^2/C_l^{TT}} \end{bmatrix} \begin{bmatrix} n_{lm}^T \\ n_{lm}^E \end{bmatrix} = \begin{bmatrix} a_{lm}^T \\ a_{lm}^E, \end{bmatrix} \tag{76}$$

where n_{lm}^T and n_{lm}^E are unit norm, complex random numbers. Since a_{lm}^T are known constraints, we can solve the system of equations:

$$\sqrt{C_l^{TT}} n_{lm}^T = a_{lm}^T$$
$$C_l^{TE}/\sqrt{C_l^{TT}} n_{lm}^T + \sqrt{C_l^{EE} - (C_l^{TE})^2/C_l^{TT}} n_{lm}^E = a_{lm}^E, \tag{77}$$

substituting the first equation into the second to arrive at an expression for a_{lm}^E in terms of a_{lm}^T and the power spectra, C_l. To ensure the maps are real, the condition $a_{lm}^* = (-1^m)a_{l-m}$ is demanded.

Naively, we might expect the constrained a_{lm}^E to have lower variance in its power spectrum than the unconstrained a_{lm}^E since there is significant *TE* correlation and the temperature component is fixed. However, this is not necessarily the case. When the a_{lm}^T fluctuate high, the mean of the constrained ensemble of a_{lm}^E's is also high, resulting in increased variance in the ensemble of simulations. For particularly high a_{lm}^T, this can result in larger variance for the constrained simulations than the unconstrained simulations. This happens to be the case in the BICEP field near $l = 150$, as seen in Figure 23.

In practice, we take the a_{lm}^T from the *Planck* temperature Needlet Internal Linear Combination (NILC) map (Planck Collaboration XII 2014). This map uses the multi-frequency coverage of *Planck* to remove the galactic contribution to microwave emission, leaving a high signal to noise map of the CMB temperature field. The map has some contamination near the galactic plane. However, we have found that the impact of this is very local and does not affect the higher galactic latitudes where the BICEP field is located. The noise level in the *Planck* temperature map is fractionally small compared to the temperature signal, and we have found that this noise contributes a similarly small fraction to the constrained realization of *E*-modes.

A.2. Lensing

The temperature anisotropies in the *Planck* NILC map have been lensed by the intervening structure between us and the surface of last scattering. This means the a_{lm}^T calculated from the *Planck* NILC map contain the effects of lensing and when used in Equation (77), the lensing distortion propagates through to the constrained a_{lm}^E. Ideally, the a_{lm}^T in Equation (77) would be from the unlensed sky, however, in the absence of an accurate map of the lensing deflection field, we have no way of de-lensing the a_{lm}^T.

Because our power spectrum analysis is insensitive to the off-diagonal correlations among modes with $l = l'$ that are produced by lensing, a reasonable workaround for this problem is to use the Planck NILC but also to use the lensed C_l^{TT} spectrum. The result is an ensemble of a_{lm}^E for the lensed a_{lm}^T, but which have the correct covariance given by C_l^{TE}. For the multipole range of interest in BICEP2, lensing has a small impact on a_{lm}^T. We use the unlensed power spectra for C_l^{TE} and C_l^{EE}.

Another subtlety incorporating lensing into constrained realizations is the question of how to simulate lensing of the polarization sky. This can be accomplished using LensPix (Lewis 2011) to numerically lens the primordial *E*-modes, which creates *B*-modes with the correct statistics. For the constrained realizations, this is impossible because the NILC map is lensed by the true sky lensing field. The true sky lensing field is not known with high signal to noise, and therefore we cannot lens the *E*-modes by the same field.

Our solution to this problem is to lens the *E*-modes with a random realization of the lensing field, but use the *Planck* 143 GHz map as the temperature field. This procedure ignores lensing correlations in *TE* and *TB* because the deflection field is different for the polarization and temperature. However, the mean of the lensing *TE* and *TB* is zero, and the
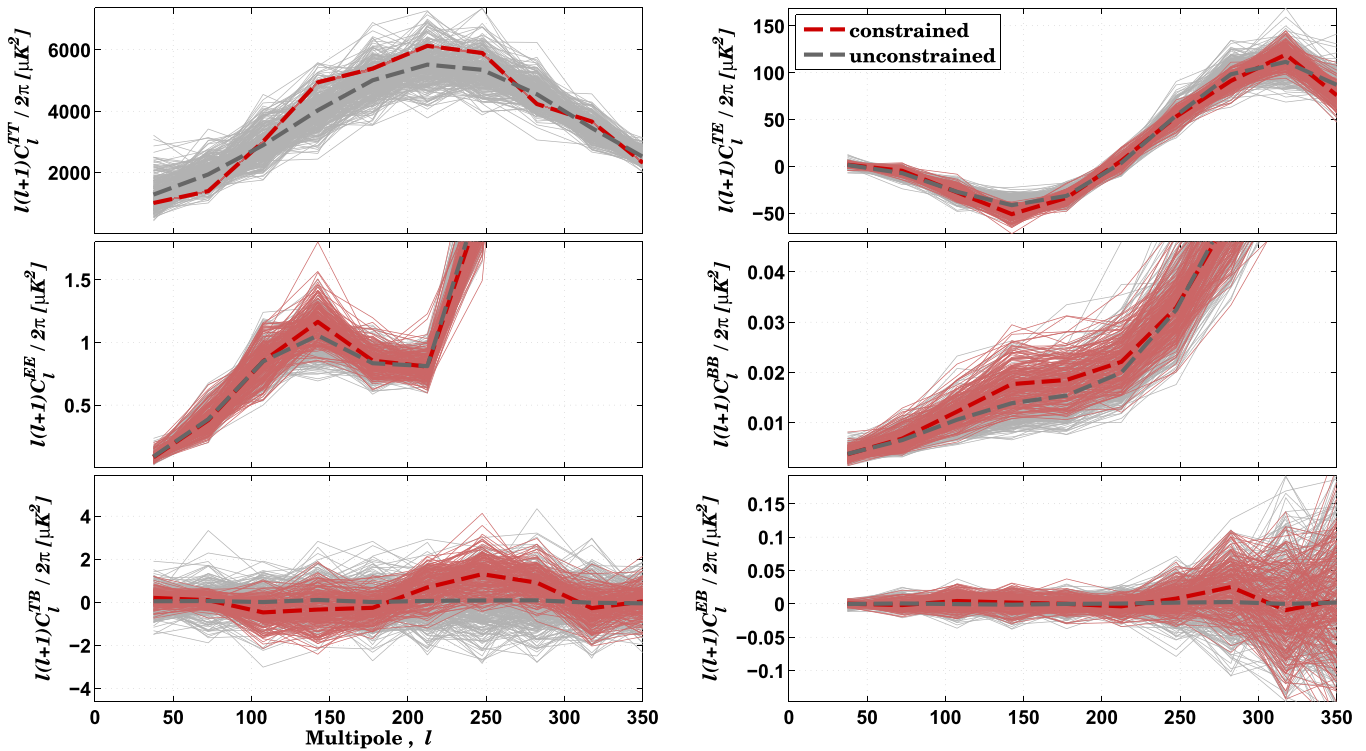
Figure 23. Power spectra of unconstrained and constrained simulations. For the constrained simulations the *T* sky is fixed to the Planck NILC map. By chance, the *TT* power in the Bicep field is above average near $l = 150$, and this leads to increased power and variance in *TE* and *EE* for these multipoles. The *BB* spectrum is computed with the Smith estimator for both constrained and unconstrained simulations in order to provide an equal comparison between the two as we cannot use the matrix estimator on unconstrained simulations. It therefore contains both E/B leakage and lensing signal. The leaked *B*-mode power creates a significant *TB* signal in the constrained simulations since the leaked *B*-modes correlate with the temperature template sky.

additional variance in *TE* and *TB* caused by lensing is small enough that it can be ignored on degree scales.

REFERENCES

Aikin, R. W. 2013, PhD thesis, California Institute of Technology
BICEP1 Collaboration 2014, ApJ, 783, 67
BICEP2 Collaboration I 2014, PhRvL, 112, 241101
BICEP2 Collaboration II 2014, ApJ, 792, 62
BICEP2 Collaboration III 2015, ApJ, 814, 110
BICEP2/*Keck* and Spider Collaborations 2015, ApJ, 812, 176
Bischoff, C., Hyatt, L., McMahon, J. J., et al. 2008, ApJ, 684, 771
Bond, J., Jaffe, A. H., & Knox, L. 1998, PhRvD, 57, 2117
Bunn, E. F., & White, M. 1997, ApJ, 480, 6
Bunn, E. F., Zaldarriaga, M., Tegmark, M., & Oliveira-Costa, A. D. 2003, PhRvD, 67, 023501
Dvorkin, C., Peiris, H. V., & Hu, W. 2008, PhRvD, 77, 063008
Goldberg, J. N., Macfarlane, A. J., Newman, E. T., Rohrlich, F., & Sudarshan, E. C. G. 1967, JMP, 8, 2155
Gorski, K., Hivon, E., Banday, A., et al. 2005, ApJ, 622, 759
Grain, J., Tristram, M., & Stompor, R. 2009, PhRvD, 79, 123515
Hamaker, J. P., & Bregman, J. D. 1996, A&AS, 117, 161
Hanson, D., Hoover, S., Crites, A., et al. 2013, PhRvL, 111, 141301
Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2

Hu, W., Hedman, M. M., & Zaldarriaga, M. 2003, PhRvD, 67, 043004
Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, PhRvD, 55, 7368
Keck Array & BICEP2 Collaborations V 2015, ApJ, 811, 126
Keck Array & BICEP2 Collaborations VI 2016, PhRvL, 116, 031302
Kovac, J. M., Leitch, E. M., Pryke, C., et al. 2002, Natur, 420, 772
Kuo, C. L., Ade, P. A. R., Bock, J. J., et al. 2004, ApJ, 600, 32
Lawson, C. L., Hanson, R. J., Kincaid, D. R., & Krogh, F. T. 1979, ACM Trans. Math. Softw, 5, 308
Lewis, A. 2011, LensPix: Fast MPI full sky transforms for HEALPix, Astrophysics Source Code Library, ascl:1102.025
Planck Collaboration I 2015, arXiv:1502.01582
Planck Collaboration XII 2014, A&A, 571, A12
Polarbear Collaboration 2014, ApJ, 794, 171
Pryke, C., Ade, P., Bock, J., et al. 2009, ApJ, 692, 1247
Readhead, A. C. S., Myers, S. T., Pearson, T. J., et al. 2004, Sci, 306, 836
Sheehy, C. D. 2013, PhD thesis, University of Chicago
Shimon, M., Keating, B., Ponthieu, N., & Hivon, E. 2008, PhRvD, 77, 083003
Smith, K. M. 2006, PhRvD, 74, 083002
Smith, K. M., & Zaldarriaga, M. 2007, PhRvD, 76, 043001
Tegmark, M. 1997, PhRvD, 55, 5895
Tegmark, M., & de Oliveira-Costa, A. 2001, PhRvD, 64, 063001
van Engelen, A., Sherwin, B. D., Sehgal, N., et al. 2015, ApJ, 808, 7
Zaldarriaga, M. 1998, ApJ, 503, 1
Zaldarriaga, M., & Seljak, U. 1997, PhRvD, 55, 1830