Limits on Polarized Dust Spectral Index Variations for CMB Foreground Analysis

Keisuke Osumi, Janet L. Weiland, Graeme E. Addison, and Charles L. Bennett

Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA; keisukeumi@gmail.com

Received 2021 April 2; revised 2021 August 10; accepted 2021 August 16; published 2021 November 12

Abstract

Using Planck polarization data, we search for and constrain spatial variations of the polarized dust foreground for cosmic microwave background (CMB) observations, specifically in its spectral index, \( \beta_d \). Failure to account for such variations will cause errors in the foreground cleaning that propagate into errors on cosmological parameter recovery from the cleaned CMB map. It is unclear how robust prior studies of the Planck data that constrained \( \beta_d \) variations are due to challenges with noise modeling, residual systematics, and priors. To clarify constraints on \( \beta_d \) and its variation, we employ two pixel space analyses of the polarized dust foreground at \( >3.7 \) scales on \( \approx 60\% \) of the sky at high Galactic latitudes. A template fitting method, which measures \( \beta_d \) over three regions of \( \approx 20\% \) of the sky, does not find significant deviations from a uniform \( \beta_d = 1.55 \), consistent with prior Planck determinations. An additional analysis in these regions, based on multifrequency fits to a dust and CMB model per pixel, puts limits on \( \sigma_{\beta_d} \), the Gaussian spatial variation in \( \beta_d \). The data support \( \sigma_{\beta_d} \) up to 0.45 at the highest latitudes, 0.30 at midlatitudes, and 0.15 at low latitudes. We also demonstrate that care must be taken when interpreting the current Planck constraints, \( \beta_d \) maps, and noise simulations. Due to residual systematics and low dust signal-to-noise ratios at high latitudes, forecasts for ongoing and future missions should include the possibility of large values of \( \sigma_{\beta_d} \) as estimated in this paper, based on current polarization data.

Unified Astronomy Thesaurus concepts: Cosmic microwave background radiation (322)

1. Introduction

Since its discovery by Penzias & Wilson (1965), the cosmic microwave background (CMB) has played a central role in cosmology. The CMB intensity and anisotropy have been well studied and have provided some of the strongest evidence and constraints for the prevailing \( \Lambda \)CDM model of the universe (e.g., Bennett et al. 2013; Planck Collaboration VI 2020e). Measuring the CMB polarization improves these constraints. In particular, the CMB polarization anisotropy can be used to constrain reionization (e.g., Page et al. 2007; Planck Collaboration VI 2020e; Pagano et al. 2020; Watts et al. 2020) and provide an indirect detection of primordial gravitational waves (e.g., Kamionkowski et al. 1997). For this reason, there are a number of existing and future experiments designed to undertake increasingly challenging measurements of the CMB polarization anisotropy.1

One of the major analysis challenges for CMB experiments is that of separating foreground and CMB components. This is especially true for polarized observations, where emission from Galactic synchrotron and thermal dust emission dominates the fainter polarized CMB signal, even at high Galactic latitudes. We focus on the polarized dust foreground, for which the Planck multifrequency high-frequency instrument (HFI) data provide the current best high-latitude sky coverage.

The observed dust signal is a superposition of emission from the dust grain population along the line of sight, which includes grains with differing sizes, shapes, and compositions subject to ambient radiation fields, all of which affect the steady-state dust temperature, \( T_d \), and the observed spectral energy distribution (SED; e.g., Draine & Lee 1984; Weingartner & Draine 2001). Aspherical dust grains tend to be aligned with the longest axis perpendicular to the local magnetic field, and their most efficient emission or absorption is along the longest axis, which produces an overall polarization (e.g., Lazarian & Finkbeiner 2003; Lazarian 2007). This alignment, and hence the polarization, also depends on the properties of the dust grains (Hildebrand et al. 2000).

Given the complexity of physics leading to the observed dust signal along a line of sight, one could attempt to fit a multicomponent model with a range of dust populations with different dust temperatures, \( T_d \), and spectral indices, \( \beta_d \). A priori, it would not be surprising if \( \beta_d \) varies significantly across the sky as different dust populations with different compositions of dust grains are considered. However, Planck measurements indicate a large degree of uniformity of the dust foreground across most of the sky. The Planck Commander pixel-based analysis for public data release 3 (PR3) found a value of \( \beta_d = 1.6 \pm 0.13 \) over the sky (Planck Collaboration IV 2020d). Similar results were found for the PR4 foreground analyses (NPIPE, Planck Collaboration Int. LVII 2020c), based on these findings, Planck Collaboration IV (2020d) and Planck Collaboration Int. LVII (2020d) adopted a Gaussian prior of \( \beta_d = 1.6 \pm 0.1 \) for the fit in their final foreground model. However, when a Gaussian prior is omitted from the fit, large tails are present in the recovered \( \beta_d \) distributions in Figure 29 of Planck Collaboration IV (2020d). These were ascribed to poor fits in noise-dominated regions of the sky but could also be due to spatial variation in \( \beta_d \).

A more focused study of the dust foreground was conducted on the older Planck PR2 data set in Planck Collaboration Int. XXII (2015). This study considered \( \beta_d \) on disjoint 10° circular patches but treated each disjoint patch as an independent measurement of a spatially uniform \( \beta_d \) and did not attempt to characterize spatial variations in \( \beta_d \).

Planck Collaboration XI (2020d) also conducted a focused study of the dust foreground for the Planck PR3 data set, which significantly reduced the level of residual systematics. They computed the angular power spectra in six overlapping nested regions with effective sky fractions of 24%–71% to determine a

---

1 A list of experiments and their properties is provided at https://lambda.gsfc.nasa.gov/product/expt/.
constraint in the allowed Gaussian variation in $\beta_d$ of $\sigma_{\beta_d} = 0.092 \pm 0.005$. However, this constraint is defined as the standard deviation over the six highly correlated $\beta_d$ estimates from their six nested regions, which limits its applicability.

Spatial variations in $\beta_d$ can also produce a decorrelation of the B-mode power spectrum between the Planck 217 and 353 GHz polarization data, which is expected to be more significant at smaller ($\ell > 50$ or $< 3^{\circ}7$) scales than we investigate in this work. This was explored in Planck Collaboration Int. L (2017), Sheehy & Slosar (2018), and Planck Collaboration XI (2020a). Sheehy & Slosar (2018) and the later Planck analysis in Planck Collaboration XI (2020b) did not find a significant decorrelation in the B-modes. Upon including decorrelation in their CMB analysis, BICEP2 Collaboration et al. (2018) found that their constraints were not significantly shifted and little evidence for decorrelation, at least given the current experimental noise.

In the analyses cited above, it is possible to achieve stronger constraints on $\beta_d$ and its spatial variation by introducing intensity data, since the dust intensity has a higher signal-to-noise ratio. Although Planck Collaboration XI (2020a) found that the mean sky $\beta_d$ was consistent between polarization and intensity, differences can arise, for example, by magnetic dipole emission due to inclusions randomly oriented within interstellar dust grains, which can reduce the polarization of the dust emission (Draine & Hensley 2013).

Several studies have shown evidence for spatial variation in $\beta_d$. Recently, Pelgrims et al. (2021) found that combining Planck polarization data with neutral hydrogen measurements and leveraging a frequency decorrelation analysis can show detectable differences in the dust SED along specific lines of sight targeted by H1 observations. As mentioned earlier, other analyses conducted over large sky regions also show hints of spatial variation in $\beta_d$ but include complicating factors beyond instrumental noise, like the choice of noise model, prior choices, and residual systematics (e.g., Planck Collaboration Int. XXII 2015; Planck Collaboration X 2016; Planck Collaboration Int. L 2017; Planck Collaboration IV 2020a; Planck Collaboration XI 2020b; Planck Collaboration Int. LVII 2020c). These additional features are difficult to disentangle from true on-sky $\beta_d$ variations.

To shed some light on this question, we engage in a pixel space investigation of the power the Planck polarization data have to constrain $\beta_d$ variations on $\approx 60\%$ of the sky at high Galactic latitudes. We chose to work in pixel space because both the dust foreground signal and the primary challenges we face in this analysis (residual systematics and the accuracy of the noise model) have large-scale spatial structures that are easier to account for in pixel space. The power spectra of foreground signals lose information, and the data compression it provides is not necessary in this analysis, as we restrict ourselves to large angular scales at low resolution ($>3^{\circ}7$).

This paper is structured as follows. Section 2 specifies the Planck data products used. Section 3 describes the foreground model and how the realistic simulations were constructed. Section 4 presents the first analysis method, which is based on a two-band template fit, with corresponding results included in Section 5. Section 6 describes the more flexible second analysis we conducted, which was based on a multifrequency fit of the model to the data, with a description of results in Section 7. These include a comparison between the simulations and the data, with findings on the power of the Planck polarization data to constrain $\beta_d$ variations. Finally, we conclude in Section 8.

We provide a road map to this paper in Table 1, which may be a useful guide to the reader.

## 2. Data and Masks

### 2.1. Data

In this work, we analyze the 2018 PR3 Planck Stokes $Q$ and $U$ maps at the four highest polarized frequency channels from the Planck HFI at 100, 143, 217, and 353 GHz using applicable correction maps and their respective noise products, which include variance maps and noise simulations (Planck Collaboration III 2020b). We use the Planck 30 GHz polarization $Q$ and $U$ maps as a template to remove the synchrotron emission present in the HFI frequencies and exclude the low-frequency instrument (LFI) 44 and 70 GHz channels from the analysis. As discussed in Section 7.1, LFI data provide little constraining power on the synchrotron spectral index at the pixel scales used here. Use of the template removal approach for synchrotron therefore has two advantages. First, it avoids the effects of the poorly understood noise properties at 44 GHz (e.g., BeyondPlanck Collaboration 2020). Second, it mitigates the potential

| Sec. | Figure/Table | Analysis | $\beta_d$ | Bands |
|------|--------------|----------|-----------|--------|
| 5    | Table 2      | Dust template fit after CMB removal over 3 disjoint regions | Solved | 217 and 353 GHz |
| 7.1  | Figure 3     | K-S tests for $\chi^2$ PTE consistency between data and simulations | 1.55 fixed | 100–353 GHz |
|      | Tables 3 and 4|          |           |        |
| 7.2  | Figure 7     | Residual 353 GHz dust correlation | 1.35–1.75 stepped | 100–353 GHz |
|      | Table 4      | PTE consistency between data and simulations | 1.55 fixed |        |
| 7.3  | Figure 8     | Injecting random Gaussian variation | Gaussian mean = 1.55 | 100–353 GHz |
|      | Table 4      |          |           |        |
| 7.4  | Figure 10    | Injecting GNILC or Commander variations | Fixed $\beta_d$ and $T_d$ from GNILC or Commander | 100–353 GHz |

### Table 1

Road Map of Analyses in This Paper
introduction of spurious large-scale power into the analyses from calibration systematics in the LFI bands (Weiland et al. 2018; Planck Collaboration II 2020a).

Since the release of the PR3 products, there have been two new data releases using alternative mapmaking pipelines, SRoll2 (Delouis et al. 2019) and NPIPE (Planck Collaboration Int. LVII 2020c). The analyses in this paper rely in part on estimates of the contribution of the CMB to the dust component separation, which are dominated by data in the HFI frequency bands. The Planck Collaboration provides a number of CMB component estimates based on the PR3 maps (see Section 2.2). Equivalent products do not accompany the SRoll2 release, and there is a substantial CMB mapping transfer function associated with the NPIPE maps. Although there are also processing improvements associated with the SRoll2 and NPIPE products, we chose to proceed with PR3.

As is common in studies of CMB foregrounds in the pixel domain, the analysis requires the maps at each frequency considered to be at a common resolution, as defined by the smoothing scale and the HEALPix (Górski et al. 2005) $N_{\text{side}}$ pixel resolution parameter. Therefore, before conducting the foreground analysis, we smooth each Planck data product to a common resolution of 2° FWHM and downgrade as necessary to $N_{\text{side}} = 16$.

The foreground analysis also requires a noise model per pixel per frequency, for which the default variance maps packaged with the $Q$ and $U$ maps are insufficient. Although the Planck processing pipeline removes most instrumental effects from the data, there are residual systematics that are not removed or accounted for in the default variance maps. These residual systematics are modeled in Planck’s Full Focal Plane 10 (FFP10) noise simulations (Planck Collaboration III 2020b). These are end-to-end simulations that include all significant systematic effects at the time-ordered-data level, e.g., the analog-to-digital converter nonlinearity and bolometer time response. However, due to computational constraints, the FFP10 simulations omit several components, like the effect of discrete point sources, glitching/deglitching, and far side lobes. There are 300 such simulations available at each frequency.

We leveraged these 300 simulations per frequency to estimate variance maps that accounted for known residual systematic effects. This was done by replicating the treatment of the $Q$ and $U$ data maps on the FFP10 simulated noise $Q$ and $U$ maps to produce 300 FFP10 noise simulations at the analysis resolution. By computing variances per pixel over these 300 downgraded FFP10 simulations, we estimated the $QQ$, $UU$, and $QU$ variance maps required for the analysis.

These estimated variance maps alone do not capture deviations from zero-mean Gaussian noise in the data. To account for such deviations, rather than generating new Gaussian noise maps, we use the FFP10 noise simulations as the noise component in the suite of simulations we calibrate the analysis against. This procedure is described in Section 3.2.

### 2.2. Additional Planck Products

Portions of the analysis also make use of CMB and dust foreground products provided in Planck public releases. In particular, we utilize products from four different foreground analyses conducted by the Planck Collaboration: Commander, Generalized Needlet Internal Linear Combination (GNILC), Spectral Estimation Via Expectation Maximization (SEVEM), and Spectral Matching Independent Component Analysis (SMICA). We use the CMB estimates from all four analyses and, since only Commander and GNILC produce a dust model, the dust model from these two analyses.

The Commander approach fits a parametric model to the data using a Gibbs sampling Monte Carlo method (Eriksen et al. 2004; Seljebotn et al. 2017). This model includes terms for the CMB and foregrounds and can be performed in the map or angular power spectrum domain. We use the Commander CMB and dust products packaged with the earlier PR2 data release (Planck Collaboration I 2016), as the publicly available PR3 Commander products do not include maps of the dust temperature, which are necessary to compare the foreground results against.

The GNILC approach extracts maps for each component in the data by computing the minimum-variance linear combination of input maps in a needlet spherical wavelet basis (Remazeilles et al. 2011). The GNILC extends the classic internal linear combination method (Bennett et al. 2003), which can only extract a component that is spatially uniform and rigidly scales with frequency, to account for more complex spatial and spectral behavior that allows for fits for foreground models. As the PR3 GNILC products do not include maps of the $T_d$ or $\beta_d$, we revert to the earlier PR2 GNILC CMB and dust products for these quantities.

The SEVEM approach uses templates to separate the CMB from the foregrounds (Martínez-González et al. 2003; Leach et al. 2008; Fernández-Cobos et al. 2012). Foreground templates are generated using the highest- and lowest-frequency bands and used to remove the foreground from the CMB-dominated intermediate bands. As a consequence, SEVEM does not output a dust model, but we make use of the PR3 SEVEM CMB maps.

SMICA (Cardoso et al. 2008) is a nonparametric method that fits the amplitude and spectral parameters of the CMB and foregrounds in harmonic space. These values are used to derive harmonic space weights that are used to combine the data maps to produce a CMB map. As with SEVEM, we use the PR3 SMICA CMB maps in part of the analysis.

### 2.3. Masks

We define three disjoint sky regions that we treat independently in the analyses. These regions are chosen to probe different dust signal-to-noise ratio regimes while excluding regions with complicated foreground structure and residual systematics along the Galactic plane. Each of these three regions covers approximately 20% of the sky for a total analyzed sky fraction of about 0.6.

Figure 1 depicts the masks used to define the disjoint regions we consider in this work. These masks are based on downgraded versions of those provided by Planck and used in Planck Collaboration X (2016). The original Planck masks were produced by thresholding on the estimated total foreground amplitude in a pixel target certain sky fractions. The masks are produced by taking the difference between these Planck masks from consecutive sky fractions.

### 3. Model and Simulations

#### 3.1. Model

The $Q$ and $U$ data in each pixel for each frequency, $m(\nu, l_{Q,U})$, can be modeled as a sum of the CMB, noise, and dominant foregrounds in polarization: synchrotron and thermal dust. This model, similar to that adopted by Planck (Planck...
Collaboration IV 2020d), represents the synchrotron as a power law in frequency and the dust as a modified blackbody and holds the CMB component fixed across frequency, as the CMB is constant in the thermodynamic temperature units of the data:

\[ m(\nu)_{Q,U} = m_{\text{sync}}(\nu)_{Q,U} + m_{\text{dust}}(\nu)_{Q,U} + \text{CMB}_{Q,U} \]

\[ = A_{s,\nu} \frac{g_{s}(\nu)}{g_{e}(\nu)} \left( \frac{\nu}{\nu_{s}} \right)^{\beta_{s}} \]

\[ + A_{d,\nu} \frac{g_{d}(\nu)}{g_{e}(\nu)} \left( \frac{\nu}{\nu_{d}} \right)^{\beta_{d}+1} \frac{\exp\left(\frac{h\nu}{k_{B}T_{d}}\right) - 1}{\exp\left(\frac{h\nu}{k_{B}T_{d}}\right)} - 1 \]

\[ + \text{CMB}_{Q,U}. \]  

(1)

Here the subscript \([Q, U]\) denotes a two-component vector for \(Q\) and \(U\), \(\nu\) is the frequency, \(A_{s}\) is the synchrotron foreground amplitude, \(A_{d}\) is the dust foreground amplitude, \(\nu_{s}\) is the synchrotron reference frequency (30 GHz), \(\nu_{d}\) is the dust reference frequency (353 GHz), \(\beta_{s}\) is the synchrotron spectral index, \(\beta_{d}\) is the dust spectral index, \(T_{d}\) is the dust temperature, \(h\) is Planck’s constant, and \(k_{B}\) is the Boltzmann constant. The factors of \(g_{e}(\nu)\), where \(g_{e}\) refers to synchrotron (s) or dust (d), convert the brightness temperature units of the foreground models to the thermodynamic temperature units of the data.

Bandpass-integrated color corrections, \(C_{\nu}(\nu)\), are needed to refer the true spectrum to the adopted reference spectral dependence of the data. For synchrotron, we interpolate for \(\beta_{s} = -3.1\) from a grid of correction values provided in the Planck Explanatory Supplement.\(^2\) For dust, we use a UC_CC IDL code derived in Planck Collaboration IX (2014) and provided by the Planck Collaboration\(^3\) to calculate \(C_{\nu}(\nu)\) for a grid of varying \(\beta_{d}\) and \(T_{d}\) values that are interpolated over for each \(\beta_{d}\) and \(T_{d}\) value used in the fits. Additionally, conversion factors \(U(\nu)\) are needed to convert from the brightness temperature units of the foreground model to the thermodynamic temperature units of the data. We use the values for \(U(\nu)\) provided for each frequency in Table 2 of Planck Collaboration XI (2020f), which were also derived using the same UC_CC IDL code.

Given \(C_{\nu}(\nu)\) and \(U(\nu)\), the \(g_{\nu}(\nu)\) factors in Equation (1) above are defined as

\[ g_{\nu}(\nu) = \frac{C_{\nu}(\nu)}{U(\nu)}. \]  

(2)

3.2. Simulations

These analyses must be calibrated against accurate simulations of the Planck data, which are known to contain non-Gaussianities and have a nonzero noise mean (Planck Collaboration III 2020b).

These simulations are generated at \(N_{\text{sims}} = 16\) and defined as a sum of foregrounds, CMB, and noise. This is the model as described in Equation (1), but with an additional noise term.

The foreground component was constructed by taking the Planck 30 GHz map as the synchrotron amplitude and the 353 GHz map as the dust amplitude and scaling to all intermediate frequencies assuming the Planck best-fit sky averages for \(\beta_{s} = -3.1\), \(\beta_{d} = 1.55\), and \(T_{d} = 19.6\) K (Planck Collaboration IV 2020d; Planck Collaboration XI 2020f).

The CMB component was generated with a fiducial input cosmology, with \(10^{5}A_{s}\) e\(^{-2}\) = 1.881, \(\tau = 0.06\), and all other cosmological parameters set to those of the 2015P(CHM) (Planck) cosmology in Table 6 of Planck Collaboration XIII (2016). The noise component was set to be one of the downgraded FFP10 Planck noise simulations rather than a Gaussian noise realization based on the noise variance estimates. This allows us to take advantage of the Planck Collaboration’s systematics modeling. As there are only 300 FFP10 noise simulations available, we can produce only 300 baseline simulations.

Here we have described the baseline simulations, but we also extend these simulations in Sections 7.3 and 7.4 by introducing variations in the dust foreground to investigate their effects on the analyses.

4. Two-band Template Fit

We use a two-band template fit method as a relatively simple method of probing \(\beta_{d}\) variations across large portions of the sky. This method has been used in prior Planck foreground analyses to provide independent measures of the foreground spectral indices like \(\beta_{d}\) to compare against their more involved foreground algorithms (Planck Collaboration XI 2020f).

Although this template fit method is less flexible than the more complex multifrequency fitting analysis we also implement, it directly measures \(\beta_{d}\) in different regions of the sky and helps set expectations for the level of signals we expect to see.

In the template fit method, the scaling of the dust from the template’s frequency to a map at some other frequency can be converted to a measure of \(\beta_{d}\) between the two frequencies. The template fit method can become complicated as the number of components in the maps increases, since each component can scale differently between the template and the map under consideration. However, at the highest frequencies we consider, 353 and 217 GHz, the maps are dominated by dust and, to a lesser extent, the CMB. By subtracting an estimate of the CMB beforehand, we are left with maps dominated by dust emission, which can be used to derive \(\beta_{d}\) between 353 and 217 GHz.
We perform this analysis via minimizing the following $\chi^2$ in the $Q$ and $U$ maps for the scaling factor, $\alpha$, and offsets, $a$ and $b$:

$$\chi^2 = \sum_k (Q_k - \text{CMB}_k) - \alpha Q_k - (a)^2 + (U_k - \text{CMB}_k) - \alpha U_k - (b)^2.$$  

The template scaling factor $\alpha$ can be converted to $\beta_d$ by setting $\nu_d = 353$ and $\nu = 217$ GHz in the following relation, derived from Equation (1),

$$\alpha = \frac{g_d(\nu)}{g_d(\nu_d)} \frac{\nu + 1}{\nu_d + 1} \exp\left(\frac{h\nu_d}{kT_d}\right) - 1 \exp\left(\frac{h\nu}{kT_d}\right) - 1$$

and solving for $\beta_d$.

We restrict the above analysis to the three large and disjoint regions of the sky, as defined in Section 2.3 and Figure 1, to maximize the signal-to-noise ratio and determine if there is a significant variation in $\beta_d$ between these regions.

5. Template Fit Results

The template fit analysis requires the CMB to be removed from the frequency maps beforehand. As described in Section 2.2, we use the four independently derived CMB map estimates provided in the PR3 data release: those from the Commander, SEVEM, SMICA, and GNILC algorithms.

Table 2 contains the results of the template fit analysis conducted over each of the three masked regions for each of the four CMB maps. To assign uncertainties to the $\beta_d$ estimates, we repeat the template fit analysis on each of 300 baseline simulations, assuming perfect CMB removal, and take the standard deviation of $\beta_d$ over the simulations.

We find that across regions and CMB versions, these recovered $\beta_d$ values are consistent. These values are also consistent with the overall mean and standard error $\beta_d$ value of $1.55 \pm 0.05$ recovered by Planck over the full sky (Planck Collaboration IV 2020d).

These results also suggest that it will be difficult to directly measure a variation in $\beta_d$ on the sky using the current data. The template fit analysis should let us discern large-scale region spectral differences if there are significant variations between sky regions. Further, switching the CMB estimate used has a comparable or even greater impact on $\beta_d$ values than switching the sky region. We also checked the impact of fixing the offsets, $a$ and $b$, to zero. Doing so shifts the derived $\beta_d$ by at most 0.01, or less than 0.5$\sigma$ in the strictest case in mask L, and less than the shift due to switching the CMB estimate.

We also repeated this analysis on coarser two-way partitions of the sky, such as a north-and-south split and an east-and-west split, while keeping the plane masked. No significant variations in $\beta_d$ were found.

Based on the results of this template fit analysis, rather than attempting to detect $\beta_d$ variations on the sky, we focus on the more complex multifrequency fitting analysis on extracting an upper bound on $\beta_d$ variations and testing candidate $\beta_d$ models.

6. Multifrequency Fitting

The template fit method is useful for directly computing $\beta_d$ on predetermined regions of the sky; however, it cannot detect or constrain variations in $\beta_d$ if lines of sight with different indices are assigned to the same region. This can occur due to suboptimally drawing the borders of the analyzed regions or because the scale of $\beta_d$ variations is smaller than the size of the regions we consider.

Further, the template fit method also depends on the accuracy of the CMB signal removed from the 353 and 217 GHz maps. As we discuss below in Section 6.2, the CMB maps provided by Planck contain features related to residual systematics. This may negatively affect the template fit analysis by introducing spurious correlations that scale in frequency like the CMB. Chance correlations between the dust and residual systematic signatures could shift the mean value of $\beta_d$. They could also decrease the variance, but it is more likely that the presence of residual systematics would be uncorrelated and increase the $\beta_d$ variance.

To address the concerns above, we implement a pixel-by-pixel multifrequency fitting analysis where we directly fit the parameters in Equation (1). This multifrequency fitting analysis is more flexible than the template fit analysis and allows for a test of the fidelity of the FFP10 noise simulations and the models at lower frequencies. Lower frequencies are important for cosmology, since they have higher CMB signal-to-noise ratios than 217 and 353 GHz. Multifrequency fitting also allows us to introduce both random and systematic dust variations into the simulations to validate the methodology.

6.1. Likelihood

Given the number of bands available, it is possible to leave all or most of the parameters free in Equation (1) when fitting to the data. However, there are known complications with fitting the spectral indices directly, as they exhibit non-Gaussian behavior that requires a careful treatment using priors (Dunkley et al. 2009; Planck Collaboration Int. LVII 2020c). The $T_d$ parameter is often found to be somewhat degenerate with $\beta_d$ because of the formulation of the dust SED function. Further, as we are working primarily with the HFI bands, we do not expect to constrain $\beta_d$ or $A_{d,Q}$ and $A_{d,U}$. Therefore, we perform simplified fits where we fix $\beta_s$, $\beta_d$, and $T_d$ to their best-fit sky averages as determined by the Planck Collaboration. We fix $\beta_s$ to $-3.1$ and $T_d$ to 19.6 K as in Planck Collaboration XI (2020). We hold $A_{d,Q}$ and $A_{d,U}$ fixed to the 30 GHz map we use as the synchrotron template. In our baseline fits, we also fix $\beta_d$ to 1.55 based on the earlier analysis in Section 5 and the constraints from Planck Collaboration IV (2020d). We can assess the impact of spatial variations in $\beta_d$ by changing this fixed value or the input $\beta_d$ in the simulated maps and repeating the fit.

By holding the above parameters constant within each fit, and if the noise is Gaussian, the posterior distributions for the four remaining parameters, $A_{d,Q}$, $A_{d,U}$, CMB$_Q$, and CMB$_U$, will also be Gaussian. These simplifications are justified, as even if
\( \beta_d \) and \( T_d \) are held fixed in any given fit, we can consider the impact of changing the assumed-per-fit or simulated value for \( \beta_d \) on specific dust variation statistics, defined in Section 7, to investigate the spatial variations in the dust foreground.

For each pixel, the free parameters were fit by minimizing the negative log-posterior, which we define as an effective \( \chi^2 \), or \( \chi^2_{\text{eff}} \), per pixel:

\[
\chi^2_{\text{eff}} = \sum_{\nu}(d(\nu)_{Q,U} - m(\nu)_{Q,U}) \\
\times C^{-1}(d(\nu)_{Q,U} - m(\nu)_{Q,U})^T + \sum_{\nu \in \text{Priors}} \ln(P).
\]

Here the first summation is over the four HFI frequency bands in the fit and corresponds to the usual \( \chi^2 \) with \( d \) as the data, \( m \) as the model as defined in Equation (1), and \( C \) as the QU covariance matrix for the noise in the pixel. The second summation represents the prior on CMB QU, and CMBU, which is described below in Section 6.2. Each quantity in Equation (5) is evaluated independently per pixel.

6.2. Priors

Although the simplifications described above resolve the most complicated fitting issues when defining priors, we find that applying flat priors to all of the remaining free parameters produces CMB map fits that contain recognizable systematics. To address this, we apply a Gaussian prior on the CMB pixel amplitudes that improves the recovery of the CMB in the fits.

The Gaussian CMB prior is defined by computing the covariance matrix in harmonic space for the same set of cosmological parameters used to generate the CMB for the simulations, with \( 10^2A_{s}\text{e}^{-2\tau} = 1.881, \tau = 0.06, \) and all other parameters set to those of the 2015F(J) (Plik) cosmology in Table 6 of Planck Collaboration XIII (2016). The harmonic space CMB covariance was converted to pixel space via a spherical harmonic transform operator. Although the full CMB pixel space covariance matrix is singular, it suffices for defining the CMB prior, as the fits are conducted 1 pixel at a time.

Figure 2 shows examples of the CMB map recovered from the data by Commander, which does not use a CMB prior, and the analysis in this paper, with and without this CMB prior. We find that without the CMB prior, there are large excursions from the input CMB realization that appear to be known systematic effects that trace the Planck scan pattern. These systematic patterns present in the Planck data and CMB maps must be accounted for in any analysis. We minimize such systematic excursions by applying a Gaussian prior to improve the performance of the foreground fit. This means that residual systematics are primarily confined within the foreground parameters, which translates to a high dependence on the fidelity of the noise model used in the likelihood and the simulations we calibrate against.

Use of the CMB prior differentiates the multifrequency analysis from that of, e.g., Commander. With the methodology established in this section, we proceed to implement it and show the results in Section 7.

7. Multifrequency Fitting Results

The ultimate objective is to detect or constrain possible dust variations over the sky through probes of the spatial variation of \( \beta_d \). This is straightforward for the two-band template fit analysis described above in Section 5, as we solve for a constant \( \beta_d \) over defined sky regions, but it is more complicated for the multifrequency fitting analysis we now consider. Here reliance on statistical properties inferred from multifrequency sky simulations establishes the dust variation constraints.

Constraints on variations in \( \beta_d \) over the sky are highly sensitive to the accuracy of the noise model and possible deviations from it in the form of genuine foreground variations or unaccounted-for systematics in the data that are not included in the simulations. In Section 7.1, we investigate how well the full-sky Planck-like simulations, in which we specify the CMB and foreground components that we add to the FFP10 Planck noise and residual systematic simulations, capture the behavior of the data.

We present the results of the investigations of the dust properties in the data in Sections 7.2, 7.3, and 7.4. In these sections, we make use of two statistical methods to capture the signal due to a spatially varying dust foreground.

1. We examine the total \( \chi^2 \) distribution and the \( \chi^2 \) distributions per mask region per band. A spatial variation in \( \beta_d \) should shift the \( \chi^2 \) distribution relative to the simulations where \( \beta_d \) is constant.

2. We estimate the dust contribution to the residual in each band by correlating the residual maps of the fit against the 353 GHz map we use as the dust template. If there is spatial variation in \( \beta_d \) and the fit assumes a spatially uniform \( \beta_d \), the residual maps will contain an excess dust contribution, which will appear as a nonzero correlation with the dust template. However, additional spurious correlations between the dust template and fit residual maps may be introduced as a result of the fitting analysis. Therefore, we use simulations to estimate this spurious fitting-induced uncertainty.

In Sections 7.3 and 7.4, we also introduce simulations in which the foreground dust model has an intrinsic spatial variation.

7.1. Comparisons between Data and Simulations

In this section, we work with the 300 baseline simulations generated using the prescription described in Section 3.2. These simulations represent the baseline case where the foregrounds, noise, and systematics are known. Any mismatch between these fiducial simulations and the actual data must be the result of either more complex foregrounds or unaccounted-for systematics in the data that are not included in the FFP10 simulations.

The \( \chi^2 \) distributions of the foreground fits performed on the data and the ensemble of simulations are one way of comparing the foreground fits for the simulations to those for the data. The foreground fits produce \( \chi^2 \) values per pixel per band. From these \( \chi^2 \) values, we can compute band-by-band empirical cumulative distribution functions (CDFs) for the fits to the data and the simulations. Figure 3 shows these band-by-band empirical CDFs for the \( \chi^2 \) in each of the masked regions estimated over all 300 simulations together and for the data. In general, the data and simulations appear to agree, except for all masks at 100 GHz and mask H for 143 GHz.

To quantify the (dis)agreement between the data and simulation \( \chi^2 \) distributions in Figure 3, we estimate a numerical probability-to-exceed (PTE) based on the Kolmogorov–Smirnov (K-S) test statistic. We defined a reference \( \chi^2 \) distribution as the cumulative \( \chi^2 \) distribution over the 300 baseline
simulations in each mask region in each band. Then, we computed the K-S test statistic between the \( \chi^2 \) distribution of each simulation and the reference \( \chi^2 \) distribution. We also computed the K-S test statistic between the \( \chi^2 \) distribution of the data and of the reference distribution. By comparing the K-S statistic computed in this way for the data against 300 K-S statistics computed for the baseline simulations, we numerically estimate the PTE for the data by counting the number of simulations with K-S statistic values greater than that of the data and dividing by the total number of simulations. For example, in mask H at 100 GHz, we find nine simulations with K-S statistic values greater than that of the data, from which we compute a PTE of 9/300. The results of this procedure are shown in Table 3.

Low PTE values for the data can be seen for masks M and H at 100 GHz and mask H at 143 GHz. This result is consistent with Figure 3. Although Figure 3 suggests that the PTE value of the data for mask L at 100 GHz might be low, its PTE value (\( \approx 0.38 \)) as computed in this test is not particularly low. Of the four cases discussed here, the K-S statistic for mask L at 100 GHz is the least discrepant in Figure 3, which may explain why the PTE value is acceptable.

We further consider the low PTE values shown in Table 3 at 100 GHz by numerically estimating the PTE of the data \( \chi^2 \) per pixel and directly comparing to the \( \chi^2 \) distribution for each pixel over the 300 baseline simulations. The PTE per pixel at 100 GHz is shown in Figure 4 alongside an estimate of residual unaccounted-for systematics in the 100 GHz map. This estimate of unaccounted-for systematics was produced by subtracting the SRoll2 100 GHz map and the mean of the 300 FFP10 noise simulations at 100 GHz from the PR3 100 GHz map we use. As SRoll2 reduces the residual

---

**Figure 2.** Top row: CMB \( Q \) and \( U \) maps recovered by the Planck Commander algorithm, which does not use priors on the CMB pixel amplitudes. Middle row: CMB \( Q \) and \( U \) maps recovered by the fitting analysis described in Section 6 without priors. Bottom row: same as the middle row but with a Gaussian prior on the CMB pixel amplitudes based on the 2015F(CHM) (Planck Collaboration XIII 2016). There are large excursions in the CMB maps recovered by Commander and the no-prior analyses that trace out the Planck scan pattern. Similar features are also present in the other three Planck CMB maps generated by the Planck GNILC, SEVEM, and SMICA analyses. As can readily be seen, the use of a prior substantially reduces these systematic features in the recovered CMB maps, instead largely confining them to the foreground components of the fit.
systematics, PR3 − SRoll2 yields an estimate of residual systematics present in the PR3 map. Note that SRoll2 was not used in the rest of the analysis due to the reasons outlined in Section 2.1. We estimate the noise covariances used in the fitting analysis from the 300 FFP10 noise simulations. Thus, further removing the mean of the 300 FFP10 noise simulations at 100 GHz to compute PR3 − SRoll2 − \text{mean}_{300 \text{ sims}} yields the estimate of the residual systematics unaccounted for in the FFP10 simulations for Figure 4. In Figure 4, which is in Galactic coordinates, there are bright bands along the ecliptic plane in both maps that trace out the Planck scan pattern in the top right and bottom left. These features could explain the excess in \chi^2 and discrepant PTEs at 100 GHz in Figure 3 and Table 3.

Estimating a similar PTE per pixel at 143 GHz, we find extensive regions of low PTE in mask H, consistent with the lower-than-expected \chi^2 in Figure 3 for mask H. These problems at 100 and 143 GHz are consistent with known Planck uncertainties in simulating analog-to-digital converter nonlinearity effects in the FFP10 simulations, with a known overestimate of this effect at 143 GHz (Planck Collaboration III 2020b).

As discussed above, statistical inconsistencies in some of the regions and bands may be caused by possible unaccounted-for systematic effects and mismatches in the noise model used in

Figure 3. The \chi^2 empirical CDFs from the multifrequency fits for the data and all 300 baseline simulations together broken up into the four frequency bands and three sky regions we consider. The blue curve represents the empirical CDF from all 300 simulations. For a given value of \chi^2, the corresponding value of the curve represents the fraction of pixels with smaller \chi^2 over all 300 simulations. For example, if the blue curve was 1 at \chi^2 = 10, this means that no pixel out of all the simulations has a greater \chi^2. The gray curve and shaded region correspond to the empirical CDF for the data and a 68% confidence interval around it, estimated using the DKW inequality (Dvoretzky et al. 1956). If the data and simulations agree, the blue curve should be entirely contained within this shaded region 68% of the time. The data and simulations appear consistent in most panels, except at 100 GHz and for mask H at 143 GHz.

| Band      | PTE (No. of Sims. Out of 300) |
|-----------|-------------------------------|
|           | Mask H | Mask M | Mask L |
| 100 GHz   | 9 *    | 18     | 113    |
| 143 GHz   | 8      | 104    | 175    |
| 217 GHz   | 115    | 287    | 135    |
| 353 GHz   | 56     | 70     | 165    |

Note. * Number of simulations with a K-S statistic greater than that of the data.
of the sky regions most affected by the potential presence of analog-to-digital
Galactic coordinates and at the same resolution, the bottom plot is an estimate
estimated magnitude of residual systematics in the 100 GHz
observational data to suf

Spatial variations in the polarized synchrotron spectral index,
5

against the distributions of \( \chi^2 \) over the 300 baseline simulations. Also in
Galactic coordinates and at the same resolution, the bottom plot is an estimate
of the sky regions most affected by the potential presence of analog-to-digital
nonlinearity not fully accounted for in the FFP10 noise model (see text). The
red contours denote the edges of the three masks. Comparing the top plot with
the bottom plot, there are bright bands along the ecliptic plane that trace out the
Planck scan pattern in the top right and bottom left of both maps. These
features could explain the excess in \( \chi^2 \) and discrepant PTEs at 100 GHz in
Figure 3 and Table 3.

Figure 6 shows the result of this residual–dust correlation
procedure for the data and simulations. Here, for all bands and
masked regions, the data fall within the spread in the
simulations and have consistent PTEs.

The two methods of comparing the data and simulations
presented here can both capture mismatches between the data
and simulations but have different sensitivity to different
effects. The \( \chi^2 \) distributions are more sensitive to outliers in the
frequency maps, while the correlation-based approach is more
sensitive to spatial structure in the frequency maps. Thus, it is
not surprising that the data and simulations are more discrepant
in one comparison than the other. It appears that although there
are residual systematics in the 100 and 143 GHz frequency
maps, which cause disagreement in some regions for the band-
by-band K-S statistic–based comparison, they are not strongly
correlated with the dust and average out when looking at the
total \( \chi^2 \) distributions.

7.2. Offset in \( \beta_d \)

We now investigate any overall difference between the fit \( \beta_d \)
and the simulated or true \( \beta_d \). Working with the same 300
baseline simulations from Section 7.1, where the simulated \( \beta_d \)
is 1.55, we step through the \( \beta_d \) assumed in the dust fit from
1.35 to 1.75 in steps of 0.02.

Figure 7 shows this effect in the residual–dust correlations
for the simulations and data. In most panels, the residual–dust
correlations for the simulations are consistent with no
correlation with dust at the input \( \beta_d = 1.55 \) and shift further
from zero the further the \( \beta_d \) used in the foreground fit is from
1.55. Some of the panels, which include mask H at 100 and
143 GHz, show offsets from zero correlation at \( \beta_d = 1.55 \). We
attribute these offsets to the structure of the noise in the
FFP10 noise simulations we use in the baseline simulations and
the necessity of using the simulated 353 GHz map, which
includes the CMB and noise, in the correlations to match the
treatment of the data. Repeating this analysis with white noise
in the simulations and correlating against the simulation input
dust template at 353 GHz eliminates all such offsets from zero
correlation at \( \beta_d = 1.55 \).

A similar trend holds for the data; they prefer a value near
1.55 ± 0.02 in most panels. This matches the Planck
Collaboration’s finding of a central value of ~1.55, depending
on the choice of analysis, for the full-sky \( \beta_d \) distribution
(Planck Collaboration IV 2020d). This is also consistent with
the results of the template fit analysis in Section 5. Further, the
overall behavior of the data as we step through \( \beta_d \) is consistent
with the spread in the simulations. This shows another area of
agreement between the data and simulations and between this
analysis and prior analyses.

Considering Figure 7 further, the data (and simulations) are
more constraining with steeper changes in the correlations as
we move to lower masks and thus higher signal-to-noise ratios
for the dust. Also, 217 and 353 GHz appear more constraining
than the other two bands, which is not surprising, as they are
dominated by dust.

7.3. Gaussian Variations in \( \beta_d \)

Above, we found that the overall \( \beta_d \) derived from the data
using a multifrequency fitting analysis is consistent with
expectations. Now, as a simple model of spatial variation in \( \beta_d \)

---

*Figure 4.* The top plot, in Galactic coordinates, depicts the PTE, smoothed to
5° FWHM, as estimated in each pixel at 100 GHz by comparing the data \( \chi^2 \)
against the distributions of \( \chi^2 \) over the 300 baseline simulations. Also in
Galactic coordinates and at the same resolution, the bottom plot is an estimate
of the sky regions most affected by the potential presence of analog-to-digital
nonlinearity not fully accounted for in the FFP10 noise model (see text). The
red contours denote the edges of the three masks. Comparing the top plot with
the bottom plot, there are bright bands along the ecliptic plane that trace out the
Planck scan pattern in the top right and bottom left of both maps. These
features could explain the excess in \( \chi^2 \) and discrepant PTEs at 100 GHz in
Figure 3 and Table 3.
around 1.55, we test a random Gaussian variation and set limits on the allowed $\sigma_{\beta_d}$.

To do this, rather than modifying the $\beta_d$ assumed in the fit as before, we produce 10 sets of 300 simulations by injecting a Gaussian variation defined by $\sigma_{\beta_d}$ into the uniform $\beta_d = 1.55$ map used in the simulations. We increased $\sigma_{\beta_d}$ from zero to 0.45 in 10 steps of 0.05, which defines each set of 300 simulations. For each simulation, we ran the fitting analysis with $\beta_d = 1.55$ fixed in the fits and analyzed the resulting $\chi^2$ distributions. We do not show results from the residual–dust correlations, as they are less informative for probing $\sigma_{\beta_d}$, since injecting a random variation into the simulated dust foreground should not affect the overall correlation of the residuals and the dust signal.

To set upper bounds on the observed level of $\sigma_{\beta_d}$, we use an approach similar to the K-S–based comparisons in Section 7.1. We estimate a PTE for the data compared to each set of 300 simulations for each of the 10 $\sigma_{\beta_d}$ values. Here we expect the PTE values to increase to 1 as the injected $\sigma_{\beta_d}$ increases. This is because greater $\sigma_{\beta_d}$ variations in the simulations will increase...
Figure 7. Change in the residual–dust correlations as the assumed $\beta_d$ in the fit is stepped from 1.35 to 1.75, while the simulated $\beta_d$ is held fixed at 1.55. The values and uncertainties for the simulations are taken from the mean and standard error of the mean correlation of the simulation ensemble. The errors on the data are taken from the standard deviation of the correlations for the simulations. The gray horizontal line represents zero correlation, which is expected for perfect foreground removal. The gray vertical line is at $\beta_d = 1.55$, which is the input value for the simulations and roughly the value expected from Planck constraints on the mean $\beta_d$. In most panels, both the data and simulations are consistent with each other and $\beta_d = 1.55$. We attribute instances where the simulations are offset from zero correlation at $\beta_d = 1.55$ to the structure of the noise in the FFP10 noise simulations and the necessity of using the simulated 353 GHz map, which includes the CMB and noise, in the correlations to match the treatment of the data. Repeating this exercise with white noise simulations and correlating against the simulation input dust template at 353 GHz eliminates all offsets from zero correlation at $\beta_d = 1.55$.

Figure 8 depicts the results of this procedure. This analysis sets upper bounds of $\sigma_{\beta_d}$ with a PTE = 296/300 in mask L and $\sigma_{\beta_d} = 0.3$ with a PTE = 299/300 in mask M. Mask H is the least constraining region, and we cannot set an upper bound within the range of $\sigma_{\beta_d}$ we test, consistent with having the weakest dust signal. For the highest $\sigma_{\beta_d}$ we test in mask H, 0.45, we find a PTE of 264/300.

Planck Collaboration XI (2020f) estimates $\sigma_{\beta_d} = 0.092 \pm 0.005$ via the power spectra of six nested regions on the sky with effective sky fractions of 24%–71%. This constraint is a standard deviation over highly correlated $\beta_d$ measurements for these six nested regions, so it is natural that this $\sigma_{\beta_d}$ constraint is tighter than the one derived here. Other differences may arise due to the differences between the pixel-based analysis above, which focuses on effects at a pixel scale of about 3.7°, and a power spectrum–based approach, which considers a wide range of angular scales.

The HFI noise and systematics are important limiting factors in the estimation of $\sigma_{\beta_d}$. By downgrading the map resolution further, one might hope to achieve greater sensitivity and tighten constraints on $\sigma_{\beta_d}$, although with the loss of smaller-scale information. However, the constraints will only tighten at lower resolutions if the noise and other components in the maps integrate down faster than the signal that is due to spatial variations in $\beta_d$. 
In alternative simulations with white noise, where the signal due to spatial variations in $\beta_d$ is allowed to dominate at large scales, we confirm that the constraining power improves at lower resolutions. However, there are large-scale noise and systematic structures in the data-like Planck FFP10 simulations that dominate at lower resolutions over the Gaussian variations we inject. For this reason, we find that constraints at $N_{\text{side}} = 8$ do not tighten compared to those we derive at $N_{\text{side}} = 16$. In a similar vein, dust analyses and constraints from higher-resolution analyses like those performed at $N_{\text{side}} = 128$ by Thorne et al. (2017), which found $\sigma_{\beta_d}$ in the 0.20–0.3 range, and $N_{\text{side}} = 2048$ by Planck Collaboration IV (2020d) cannot be reliably extrapolated to, e.g., $N_{\text{side}} = 16$ by assuming a statistical $N_{\text{pix}}/N_{\text{side}}$ dependence. Each constraint must be interpreted in the context of the analysis resolution. Ultimately, the analysis presented here shows that at $N_{\text{side}} = 16$, the Planck data do not rule out large ($>0.1$) $\sigma_{\beta_d}$ values, especially at higher latitudes, where the dust signal is faint.

We next examine the impact of these constraints on upcoming high-sensitivity experiments, using LiteBIRD as an example. We generate 300 simulations at $N_{\text{side}} = 16$, where simulated Planck-like 353 GHz maps were used to clean simulated LiteBIRD-like 140 GHz maps, which is the designated cosmology band. The dust amplitude was set to the Planck 353 GHz map and scaled down to 140 GHz using either a spatially constant $\beta_d = 1.55$ or a spatially varied Gaussian $\sigma_{\beta_d} = 0.15$, 0.3, or 0.45 injected per mask regions L, M, and H. Noise was added to the 353 GHz simulated maps using the same FFP10 noise simulations used in our fitting analysis. LiteBIRD white noise was added at 140 GHz based on the level specified in Table 3 in Hazumi et al. (2020). A CMB realization was also added to each simulated map, assuming the same fiducial cosmology described in Section 3.2.

The simulated 140 GHz maps were cleaned using the simulated 353 GHz maps via the same template cleaning analysis described in Section 4, with the fit restricted to the 60% of the sky in the union of the three L, M, and H mask regions. We then subtracted the input CMB from the recovered CMB to produce residual maps. PolSpice (Chon et al. 2004) was used to compute the B-mode power spectrum of these residual maps on the union of the three mask regions. Figure 9 shows the mean and standard deviation of these residual power spectra for the 300 simulations in the spatially constant and varied $\beta_d$ cases. For reference, the figure also includes theory spectra for three values of the tensor-to-scalar ratio, $r = 0.06$, 0.04, and 0.02. The current upper limit using data from a combination of experiments, including BICEP2/Keck and Planck, is $r < 0.044$ (95% CL; BICEP2 Collaboration et al. 2018; Tristram et al. 2021).

The orange spectrum in Figure 9 is derived from template-cleaned residuals in the simulated case where the maximum allowed variation in the dust spectral index is assumed. This spectrum represents an order-of-magnitude increase in residual power in the template-cleaned 140 GHz map compared to the spatially constant case shown in blue. Although even a spatially constant $\beta_d$ is not sufficient to allow a Planck-like 353 GHz polarization map to be used in future B-mode searches, the effect of these levels of $\sigma_{\beta_d}$ should be accounted for by future experiments. Relying on dust model predictions based on the assumption that $\beta_d$ is constant or only slightly varies over the high-latitude sky may be overly optimistic.

### 7.4. Injecting External $\beta_d$ Maps

Instead of injecting random variations in $\beta_d$ per pixel, we now introduce large-scale coherent spatial variations by using $\beta_d$ and $T_d$ maps produced by the Commander and GNILC foreground separation algorithms and released by the Planck Collaboration (Planck Collaboration X 2016). These maps were released as part of the 2015 PR2 data release and take advantage of the higher signal-to-noise ratio in the intensity data to make a measurement of $\beta_d$ and $T_d$ per pixel. Although the data and simulations we use are part of the 2018 PR3 data release.

![Figure 9](https://example.com/figure9.png)

**Figure 9.** Mean and standard deviation over 300 simulations for B-mode power spectra computed from residual maps after template cleaning dust + CMB + noise simulations of LiteBIRD 140 GHz using Planck 353 GHz (see text). These spectra were computed on the union of the L, M, and H mask regions using PolSpice. Theory CMB spectra for three levels of the tensor-to-scalar ratio, $r$, are shown for reference. The bump at low $\ell$ for the residual spectra is due to the structure of the Planck noise in the 353 GHz FFP10 noise simulations.
release, similar paired $\beta_d$ and $T_d$ maps for the PR3 data set were not available.

Figure 10 shows the $\beta_d$ maps we use, downgraded to the $3^\circ$7 analysis resolution. There are significant differences in the structure and amplitude of variation between the Commander and GNILC $\beta_d$ maps. We perform a similar downgrading of the $T_d$ maps and use these external maps to define the dust component in two new sets of 300 simulations, one for the Commander model and another for the GNILC model.

To investigate whether the data prefer the baseline, Commander, or GNILC dust model, we performed two additional sets of fits, where we fixed both $\beta_d$ and $T_d$ to the input spatially varying Commander or GNILC model when fitting these two new sets of simulations. This was repeated for the data. This results in two new sets of fits for the data and simulations, one for each dust model. As in the baseline case, we consider both the K-S–based $\chi^2$ comparison and the residual–dust correlations.

Table 4 includes the PTEs for these K-S comparisons. Comparing the K-S–based PTEs for the GNILC or Commander dust models to those for the baseline case, we conclude that the K-S–based comparison cannot significantly distinguish between the baseline, GNILC, and Commander dust models.

Table 4 also includes the PTEs from the residual–dust correlation distributions for these fits. There are several instances where the Commander and GNILC simulations are inconsistent with the data fit, with the data result far off in the tails of the distribution with PTEs of 1/300 or 300/300.

This is in contrast with the PTEs for the baseline dust model with $\beta_d$ fixed to 1.55 and $T_d$ fixed to 19.6 K. With the baseline model, the data lie within the distribution from the simulations for every band and mask region. Under these circumstances, the Planck polarization data can discriminate between these dust models using the residual–dust correlations and prefer a simpler dust model than that produced by the Commander or GNILC methods.

However, this does not necessarily represent a significant difference between these two dust models and the baseline model, as we have not incorporated uncertainties on $\beta_d$ or $T_d$ in the analysis, as it is not clear how to analytically propagate them to the PTEs and computationally expensive to numerically estimate. Considering the constraints on $\sigma_{\beta_d}$ from Section 7.3, it is possible that the Commander, GNILC, and baseline dust models are consistent in the masked regions we consider, despite the clear differences in amplitude and structure in Figure 10. These results emphasize that the uncertainties on the measured $\beta_d$ are significant and must be accounted for in any analysis of the Planck foreground products.

8. Conclusions

A deeper understanding of the polarized dust foreground and its possible spatial variations is motivated by its importance as one of the two dominant polarized foregrounds, along with synchrotron emission, in studies of the polarized CMB. Although mean values of $\beta_d$ for the high-latitude sky have been quoted to high precision, previous work (e.g., Planck Collaboration IV 2020d; Planck Collaboration XI 2020f) suggests that considerable spatial variation in $\beta_d$ cannot be ruled out by the Planck data, especially at higher latitudes, where the signal-to-noise ratio of the dust foreground is low. In this work, we set out to constrain such spatial variations.

We employ two different analyses in map space at a HEALPix resolution of $N_{\text{side}} = 16$ ($\sim 3^\circ$7), a two-band template fitting method and a multifrequency fitting method,
to analyze the PR3 polarized Planck frequency maps and associated data. We divide our analysis region of about 60% of the sky into three disjoint regions at high, medium, and low Galactic latitudes (the H, M, and L mask regions), each covering about 20% of the sky (Section 2). Given the complicated nature of the noise and systematics present in the Planck data, our results are calibrated against a suite of 300 simulations based on end-to-end noise simulations provided by Planck that include their best estimate of these effects (Section 3.2).

The template fit analysis takes the two highest-frequency polarization maps at 217 and 353 GHz, which are dominated by the dust foreground, removes the Planck estimates of the CMB, then forms a dust template fit with the resulting maps to directly estimate $\beta_d$ (Section 4). We perform this analysis for our 60% sky region on the H, M, and L regions separately. By considering these large partitions, we sought to maximize the signal-to-noise ratio of possible spatial variations in $\beta_d$ at large scales. Although we test several configurations and partitions for the template fit analysis, all of the estimated $\beta_d$ values for all of our sky partitions are consistent within 1σ both with each other and with the Planck constraint of a mean $\beta_d = 1.55 \pm 0.05$ recovered over the full sky (Planck Collaboration IV2020d).

Using our more flexible analysis based on directly fitting a multifrequency foreground model of the CMB, synchrotron foreground, and dust foreground in each pixel, we find the following.

1. When fitting the CMB and foregrounds together for the polarized Planck data, systematics tracing out the Planck scan strategy leak into the CMB, including those CMB maps released by the Planck Collaboration. These must be accounted for in any analysis of the Planck CMB maps or foreground analyses, including in this paper (Section 6.2). These systematics are present even in the latest data releases by the Planck Collaboration (Planck Collaboration Int. LVII 2020c).

2. Although there are discrepancies between the data and the Planck noise simulations in some regions and bands, when taken as a whole, they are consistent over the 60% of the sky we consider with numerical PTEs of 168 sims/300 sims, 251/300, and 234/300 when compared to our simulation ensemble. The discrepancies in individual regions and bands may be driven by residual systematics in the data that are not completely modeled in the end-to-end simulations provided by Planck (Section 7.1).

3. Stepping through a central $\beta_d$ assumed when analyzing the data and simulations, we find again that the data prefer $\beta_d \approx 1.55$. The data and our simulations also respond consistently through all central values of $\beta_d$ considered between 1.35 and 1.75, highlighting the consistency of the data and the simulation suite (Section 7.2).

4. We can set an upper bound for Gaussian variations in $\beta_d$ defined by $\sigma_{\beta_d} = 0.15$ in the lowest Galactic latitude region we consider, 0.30 at midlatitudes, and essentially unconstrained at high latitudes, though we find that the data support $\sigma_{\beta_d} \approx 0.45$. These constraints are based on comparisons to 300 simulations from which we estimate numerical PTEs of 297/300, 299/300, and 264/300 at these levels of $\sigma_{\beta_d}$ (Section 7.3).

5. We apply the $\sigma_{\beta_d}$ constraints to simulations testing the quality of dust cleaning via template subtraction of Planck 353 GHz data that could achieved by future experiments, like LiteBIRD. There is an approximately order-of-magnitude increase in the B-mode residual dust spectrum at $\ell < 32$ due to these allowed levels of $\sigma_{\beta_d}$ versus a spatially constant $\sigma_{\beta_d} = 0$ (Section 7.3). This should be considered when simulating the performance of future experiments.

6. Injecting different $\beta_d$ maps derived by the Planck Collaboration into the simulations can produce significant differences in our dust variation statistics, shifting numerical PTEs comparing the data against our simulations to 1/300 or 300/300 in many cases where they were in agreement for the baseline case with a constant $\beta_d$. Taking these estimated PTEs at face value, these results suggest that the data reject these alternative spatially varying $\beta_d$ maps in favor of a spatially constant $\beta_d$. The analysis here does not consider errors on the derived $\beta_d$ maps injected into the simulations; so, the dust models tested may be entirely consistent within, e.g., the $\sigma_{\beta_d}$ bounds from above. However, these results highlight the need to account for significant uncertainties on the measured $\beta_d$ in any foreground analysis (Section 7.4).

We have not detected statistically significant spatial variations in $\beta_d$ using the polarized Planck data. Constraints on $\beta_d$ derived from the two-band template fits (Section 5) agree with those published by the Planck Collaboration. A more flexible multifrequency fitting analysis (Section 6) probes the properties of $\beta_d$ in the high-latitude regions of the sky and highlights the need for caution when interpreting all-sky mean $\beta_d$ constraints or derived $\beta_d$ maps released by the Planck Collaboration. Even though the mean $\beta_d$ value may be tightly constrained, there is considerable spatial variation allowed by the data. For example, at high latitudes, where the signal-to-noise ratio of the dust signal is low, the data support $\sigma_{\beta_d}$ at least up to 0.45.

Variations due to $\sigma_{\beta_d} = 0.45$ (high latitude), 0.30 (midlatitude), or 0.15 (low latitude) at $N_{\text{side}} = 16$ will impact cosmological parameter recovery and may be important to consider for future experiments targeting the polarized CMB. Recent simulations for quantifying the performance of future CMB facilities have assumed varying levels of $\sigma_{\beta_d}$ at various scales. For example, Simons Observatory Collaboration (2019) and The CMB-S4 Collaboration (2020) use the $d_1$, $d_4$, and $d_7$ dust models from the Python Sky Model (PySm; Thorne et al. 2017), which are all derived from the 2015 Planck Commander intensity results (Planck Collaboration X 2016) at $N_{\text{side}} = 128$. These models assume a lower variation in $\beta_d$ than the polarization data allow. Given the uncertainty in $\beta_d$, measurements from the current Planck polarization data, it may be prudent to consider more conservative dust models when optimizing the design of future surveys, like the $d_2$ ($d_3$) model in PySm, that set $\sigma_{\beta_d} = 0.25$ (0.3) at degree scales.

We thank Duncan Watts for many helpful comments. This research was supported in part by NASA grants NNX17AF34G, 80NSSC19K0526, 80NSSC20K0445, and 80NSSC21K0638. This research has made use of NASA’s Astrophysics Data System Bibliographic Services. Some of the results in this paper have been derived using the healpy and HEALPix package. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA), part of the High Energy Astrophysics Science Archive Center (HEASARC). HEASARC/LAMBDA is a service of the Astrophysics Science Division at the Astrophysical Journal, 921:175 (15pp), 2021 November 10 Osumi et al.
NASA Goddard Space Flight Center. We also acknowledge use of the Planck Legacy Archive. Planck is an ESA science mission with instruments and contributions directly funded by ESA Member States, NASA, and Canada.

ORCID iDs
Keisuke Osumi https://orcid.org/0000-0003-2838-1880
Janet L. Weiland https://orcid.org/0000-0003-3017-3474
Graeme E. Addison https://orcid.org/0000-0002-2147-2248
Charles L. Bennett https://orcid.org/0000-0001-8839-7206

References

Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003, ApJS, 148, 97
Bennett, C. L., Larson, D., Weiland, J. L., & al, E. 2013, ApJS, 208, 20
BeyondPlanck Collaboration 2020, arXiv:2011.05609
BICEP2 Collaboration, Keck Array Collaboration, Ade, P. A. R., et al. 2018, PhRvL, 121, 221301
Cardoso, J., Le Jeune, M., Delabrouille, J., Betoule, M., & Patanchon, G. 2008, ISTSP, 2, 735
Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, MNRAS, 350, 914
Delouis, J. M., Pagano, L., Mottet, S., Puget, J. L., & Vibert, L. 2019, A&A, 629, A38
Draine, B. T., & Hensley, B. 2013, ApJ, 765, 159
Draine, B. T., & Lee, H. M. 1984, ApJ, 285, 89
Dvoretzky, A., Kiefer, J., & Wolfowitz, J. 1956, Ann. Math. Statist, 27, 642
Eriksen, H. K., O’Dwyer, I. J., Jewell, J. B., et al. 2004, ApJS, 155, 227
Fernández-Cobos, R., Vielva, P., Barreiro, R. B., & Martínez-González, E. 2012, MNRAS, 420, 2162
Fuskeland, U., Andersen, K. J., Aurlien, R., et al. 2021, A&A, 646, A69
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hazumi, M., Ade, P. A. R., Adler, A., et al. 2020, Proc. SPIE, 11443, 114432F
Hildebrand, R. H., Davidson, J. A., Dotson, J. L., et al. 2000, PASP, 112, 1215
Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, PhRvD, 55, 7368
Krachmalnicoff, N., Carretti, E., Baccigalupi, C., et al. 2018, A&A, 618, A166
Lazarian, A. 2007, JQSRT, 106, 225
Lazarian, A., & Finkbeiner, D. 2003, NewAR, 47, 1107
Leach, S. M., Cardoso, J. F., Baccigalupi, C., et al. 2008, A&A, 491, 597
Martínez-González, E., Diego, J. M., Vielva, P., & Silk, J. 2003, MNRAS, 345, 1101
Pagano, L., Delouis, J. M., Mottet, S., Puget, J. L., & Vibert, L. 2020, A&A, 635, A99
Page, L., Hinshaw, G., Komatsu, E., et al. 2007, ApJS, 170, 335
Pelgrims, V., Clark, S. E., Hensley, B. S., et al. 2021, A&A, 647, A16
Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419
Planck Collaboration I 2016, A&A, 594, A1
Planck Collaboration II 2020a, A&A, 641, A2
Planck Collaboration III 2020b, A&A, 641, A3
Planck Collaboration Int. I 2017, A&A, 599, A51
Planck Collaboration Int. LVII 2020c, A&A, 643, A42
Planck Collaboration Int. XXII 2015, A&A, 576, A107
Planck Collaboration IV 2020d, A&A, 641, A4
Planck Collaboration IX 2014, A&A, 571, A9
Planck Collaboration VI 2020e, A&A, 641, A6
Planck Collaboration X 2016, A&A, 594, A10
Planck Collaboration XI 2020f, A&A, 641, A11
Planck Collaboration XII 2016, A&A, 594, A13
Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, MNRAS, 418, 467
Seljebotn, D. S., Bærlund, T., Eriksen, H. K., Mardal, K. A., & Wehus, I. K. 2017, arXiv:1710.00621
Sheehy, C., & Slosar, A. 2018, PhRvD, 97, 043522
Simons Observatory Collaboration 2019, JCAP, 2019, 056
The CMB-S4 Collaboration 2020, arXiv:2008.12619
Thorne, B., Dunkley, J., Alonso, D., & Næss, S. 2017, MNRAS, 469, 2821
Tristram, M., Banday, A. J., Górski, K. M., et al. 2021, A&A, 647, A128
Watts, D. J., Addison, G. E., Bennett, C. L., & Weiland, J. L. 2020, ApJ, 889, 130
Weiland, J. L., Osumi, K., Addison, G. E., et al. 2018, ApJ, 863, 161
Weingartner, J. C., & Draine, B. T. 2001, ApJ, 548, 296

15