A Comparison of Maps and Power Spectra Determined from South Pole Telescope and Planck Data

We study the consistency of 150 GHz data from the South Pole Telescope (SPT) and 143 GHz data from the Planck satellite over the patch of sky covered by the SPT-SZ survey. We first visually compare the maps and find that the residuals appear consistent with noise after accounting for differences in angular resolution and filtering. We then calculate that the residuals are consistent with noise after accounting for differences in angular resolution and between two independent halves of Planck data, and (3) the cross-spectrum between SPT and Planck data. We find that the three cross-spectra are well fit (PTE = 0.30) by the null hypothesis in which both experiments have measured the same sky map up to a single free calibration parameter—i.e., we find no evidence for systematic errors in either data set. As a by-product, we improve the precision of the SPT calibration by nearly an order of magnitude, from 2.6% to 0.3% in power. Finally, we compare all three cross-spectra to the full-sky errors in either data set. As a by-product, we improve the precision of the SPT calibration by nearly an order of magnitude, from 2.6% to 0.3% in power.
1. Introduction

One of the most remarkable results of modern cosmology is that a simple six-parameter model, usually referred to as the Lambda cold dark matter (LCDM) model, can fit the full range of cosmological observations. With the precision of cosmological observables now reaching the level of a few percent, however, several small discrepancies in the inferred parameter values are attracting attention. These discrepancies show up in three places. The first is in the inferred parameter values from the cosmic microwave background (CMB) compared to some observations of the local universe. For example, the amplitude of local density fluctuations, \( \sigma_8 \), that is measured from observations of large-scale structure appears lower than the \( \sigma_8 \) value implied by CMB measurements (Planck Collaboration et al. 2016c). There is also some tension between direct measurements of the Hubble parameter \( H_0 \) and the value derived from CMB measurements (Riess et al. 2011, 2016; Planck Collaboration et al. 2014c, 2016b).

Second, there are mild (1\( \sigma \)–2\( \sigma \)) discrepancies between parameter values derived from observations of the CMB by different experiments. In particular, the best-fit parameters of the LCDM model given by Planck satellite measurements of CMB temperature and polarization power spectra (Planck Collaboration et al. 2016c) are somewhat different from those derived from earlier CMB data, whether from the Wilkinson Microwave Anisotropy Probe (WMAP) satellite (Hinshaw et al. 2013), from the combination of WMAP data and data from the South Pole Telescope (SPT; Hou et al. 2014), or from WMAP + SPT plus data from the Atacama Cosmology Telescope (ACT; Calabrese et al. 2013).

Finally, work has been done on the internal consistency of LCDM model parameter values from different subsets of the Planck data. Addison et al. (2016) recently pointed out that the matter density inferred from Planck data at \( \ell < 1000 \) is 2.5\( \sigma \) discrepant from that inferred from Planck data at \( \ell > 1000 \). In response, Planck Collaboration et al. (2017) show that after correcting for certain approximations in the Addison et al. (2016) analysis, and taking into account the fact that matter density had been singled out as the most discrepant parameter, the global discrepancy is only 1.6\( \sigma \).

Taken together, these low-level discrepancies have led some to speculate that we are seeing evidence of a potential failure of the LCDM model parameter level, for instance, in Planck Collaboration et al. (2014c, 2016b), and the two data sets have been shown to be roughly consistent, but again with some low-level discrepancies.

The strength of power spectrum and parameter comparisons using the full data sets is limited by the sample variance of the SPT data at lower \( \ell \) and Planck noise at higher \( \ell \). By limiting the comparison to CMB modes measured by both experiments, we can greatly reduce the sample variance to sharpen the consistency tests between the two data sets.

Here we compare the SPT-SZ data in the 150 GHz band and Planck full-mission data in the 143 GHz band, restricted to the SPT-SZ observing region. We calculate angular cross-spectra of the SPT and Planck maps (150 × 143) and, for comparison, the cross-spectrum of one half of SPT data with the other half (150 × 150) and the cross-spectrum of one half of Planck data with the other half (143 × 143). We calculate the difference between the SPT 150 × 150 spectrum and the other two, as well as the ratio of the SPT 150 × 150 spectrum to the others.

Using simulated observations of mock skies (including realistic noise for both experiments), we calculate the expected uncertainty in these differences and ratios, and we use this to calculate the \( \chi^2 \) and probability to exceed (PTE) this \( \chi^2 \) under the null hypothesis that there are no systematic biases between the experiments. This investigation is similar to that performed by Louis et al. (2014) on ACT and Planck data over 592 deg\(^2\) of sky and two observing bands (143/148 GHz and 217/218 GHz), the conclusion of which was that ACT and Planck measured statistically consistent CMB fluctuations over that patch of sky.

This paper is organized as follows. In Section 2 we discuss the SPT 150 GHz and Planck 143 GHz temperature maps in the 2540 deg\(^2\) patch of sky that constitutes the whole of the SPT-SZ survey region. In Section 3 we present how power spectra are calculated from the maps and the simulations generated to de-bias the data and build the covariance matrix. In Section 4 we compare the power spectra and use simulations to test the null hypothesis that the two experiments are measuring the same sky, and we discuss these results in the context of comparisons between SPT and full-sky Planck data. We present our conclusions in Section 5.

2. Data

The main goal of this work is to compare maps and power spectra from the SPT 150 GHz and Planck 143 GHz data sets within the 2540 deg\(^2\) SPT-SZ survey area. As an overview of the two data sets, Figure 1 shows a sky map of Planck HFI 143 GHz full-mission data, with the SPT-SZ survey area outlined by a solid black curve. In this section, we discuss details of the mapmaking and instrumental characteristics of each experiment.

2.1. SPT

The SPT is a 10 m telescope located at the Amundsen–Scott South Pole station. From 2008 to 2011, the first camera on the SPT, a three-band bolometer array known as the SPT-SZ camera, was used to conduct a survey of \( \approx 2500 \) deg\(^2\) of the southern sky with low Galactic dust contamination, referred to as the SPT-SZ survey. As shown in Figure 1, the survey area is a contiguous region extending from 20\(^h\) to 7\(^h\) in right ascension (R.A.) and from \(-65^\circ\) to \(-40^\circ\) in declination. The survey was conducted by observing a series of 19 areas, ranging in area from roughly 100 to 300 deg\(^2\), which together form the full survey region (see, e.g., S13, Figure 2). If not specified, in this
The SPT maps are simply binned and averaged maps of filtered data and thus are biased representations of the sky. The signal in these maps is the true sky signal convolved with the instrument beam, the effect of the TOD filtering, and the effect of binning. Beams are discussed below, and the filter transfer function is estimated through simulations, as discussed in Section 3.3.

We use the S13 estimates of the SPT 150 GHz beam transfer function $B_\ell$ and its uncertainty, and we refer the reader to S13 for more details. Briefly, the main lobe is measured using bright point sources in the survey fields, while the sidelobes are measured using observations of Jupiter. Venus observations are used to stitch the two together. The main lobe of the beam is well approximated by a 1.2 arcmin FWHM Gaussian. The beam uncertainty arises from several statistical and systematic effects, including residual atmospheric noise in the maps of Venus and Jupiter, and the weak dependence of $B_\ell$ on the choice of radius used to stitch the inner and outer beam maps.

The TOD from each detector in each observation is calibrated using the response to an internal thermal source, which is in turn tied to the brightness of the Galactic H II region RCW38. For details of this calibration, see Schaffer et al. (2011). The full-depth maps are then compared to CMB satellite data to provide the overall absolute calibration. In K11 and S13, the power spectrum of the full-depth maps was compared to the full-sky Planck estimate of the CMB power spectrum in the multipole range $650 \leq \ell \leq 1000$. For this work, we initially use the calibration determined in George et al. (2015), using the comparison of SPT power spectra to the full-sky Planck power spectrum in the multipole range $670 \leq \ell \leq 1170$; however, the map-based comparison to Planck undertaken here ends up providing a significantly more precise absolute calibration, as detailed in Section 4.2.

Figure 1. Celestial southern hemisphere of CMB data from Planck HFI 143 GHz data. The black curve outlines the SPT-SZ survey coverage. The image is oriented such that the line of $0^\circ$ right ascension extends from the center to the top of the hemisphere and right ascension increases counterclockwise.

There were approximately 200 observations of each field, with each individual observation taking roughly 2 hr. Estimates of the primary CMB temperature power spectrum from this survey, as well as the resulting cosmological interpretation, are discussed in Keisler et al. (2011, hereafter K11), S13, and Hou et al. (2014). Our analysis is identical to that of S13 up to the single-observation map step. We briefly review that part of the analysis here and refer the reader to K11 and S13 for more details.

To process time-ordered data (TOD) into maps, the TOD from each SPT detector are first filtered and multiplied by a calibration factor. The filtering steps important for this analysis are a high-pass filter and the subtraction of a common mode across all detectors in each of the six 160-element modules in the focal plane. The high-pass filter cuts off signals below a temporal frequency corresponding to an angular frequency along the scan direction of $\ell = 270$. The common-mode subtraction acts as an isotropic high-pass filter with a cutoff at roughly $\ell = 500$. In both of these filtering steps, bright point sources ($S > 50$ mJy at 150 GHz) are masked to avoid large filtering artifacts.

The TOD from individual SPT detectors are then binned into maps using inverse-variance weighting, i.e., TOD samples corresponding to times when an individual detector was pointed at a particular pixel are averaged together using inverse-variance weighting and then assigned to that pixel. Maps are made using the oblique Lambert equal-area azimuthal projection (Snyder 1987) with 1 arcmin square pixels.
numerical stability of our unbiased power spectrum calculations. In the power spectrum analysis described below, we use the projected Planck maps generated from the mapmaking pipeline with this simple high-pass filter applied to the Planck mock TOD.

The Planck 143 GHz beam has been measured using the Planck TOD around planets, such as Mars, Jupiter, and Saturn (Planck Collaboration et al. 2014b, 2016f). The beam window function $B(\ell)$ for individual frequency detector sets has been released by the Planck Collaboration as part of the “Reduced Instrument Model” (RIMO). From the 2013 Planck data release to the 2015 data release, there was a marked improvement in the beam characterization, such that the uncertainty on the 2015 beam window functions for 143 GHz data is at the 0.1% level for the $\ell$ range of interest to this work.

3. Power Spectrum

We now turn to the methodology for calculating unbiased cross-spectra of SPT 150 GHz maps and Planck 143 GHz maps. We introduce the power spectrum estimator in Section 3.1. Simulations are used in several places while estimating the power spectra; we describe how these simulations are generated in Section 3.2. We discuss how these simulations are used to estimate the SPT filter transfer function in Section 3.3 and to calculate the noise and sample variance parts of the power spectrum covariance in Section 3.4.1. We describe how beam uncertainty is incorporated into the analysis in Section 3.4.2. Finally, we discuss how to calculate the window functions needed to compare the binned cross-spectrum estimates—or “bandpowers”—to a theory spectrum in Section 3.4.3.

3.1. Power Spectrum Estimator

We use an estimator similar to that used in K11 and S13 to calculate the various cross-spectra in this work. Each cross-spectrum is calculated by correlating maps made from different observations of the same field—either correlating full-depth SPT maps with full-depth Planck maps or correlating two half-depth maps from the same experiment. In the latter case, there is a noise penalty relative to K11 or S13, who used $O(100)$ independent maps depending on the field. However, the noise penalty is largely insignificant for the SPT maps (less than 2% in power) on the angular scales of interest, and it is unavoidable for the Planck maps because a larger number of independent splits are not publicly available. We mask and zero-pad each map before calculating its two-dimensional Fourier transform, $\tilde{m}_\ell$. The mask is a product of an apodization window and a point-source mask. The maps and masks are zero-padded before Fourier-transforming such that each Fourier-space pixel has a width of $\delta_{\ell} = 5$. The raw bandpowers are the binned average of the cross-spectra between two maps $A$ and $B$ within a multipole bin $b$:

$$\hat{D}^{A\times B}_b = \left\{ \frac{\ell(\ell + 1)}{2\pi} H\ell \Re [\tilde{m}^A_{\ell} \tilde{m}^{B*}_{\ell}] \right\}$$.

For the SPT-only power spectrum, $A$ and $B$ are the two SPT half-survey maps; for the SPT and Planck cross-spectrum, $A$ and $B$ are the full-mission maps of each experiment; and for the Planck-only 143 GHz power spectrum, $A$ and $B$ are the two Planck half-mission maps. $H\ell$ in the above equation represents the two-dimensional weighting of the Fourier modes that was used by S13 to handle the anisotropic noise in the SPT maps. While this weighting is suboptimal for Planck data, we still use the same S13-derived weighting for all bandpowers to minimize the differential sample variance between cross-spectra.

Since the input maps are biased estimates of the true sky, due to effects such as the application of a mask to the maps, the raw bandpowers $\hat{D}_b$ are a biased estimate of the true bandpowers, $D_b$. The biased and unbiased estimates are related by

$$\hat{D}_b = K_{bb'} D_b$$,

where

$$K_{bb'} = P_{bb'}(M_{b'f} F_{b'f} B_{b'f}) Q_{bb'}$$.

$Q_{bb'}$ is the binning operator and $P_{bb'}$ is its reciprocal, $M_{b'f}$ is the mode-coupling matrix, $F_{b'f}$ is the filter transfer function that accounts for the signal suppressed by TOD filtering, and $B_{b'f}$ is the beam function. For details on the unbiasing procedure, see K11, S13, and Hivon et al. (2002).

The unbiased bandpowers are calculated on a field-by-field basis and then combined. We combine the bandpowers obtained from individual fields using the effective area of single fields as the weighting, as in K11 and S13.

3.2. Simulations

We use simulations both in the calculation of the bandpowers and to characterize the degree of consistency between the SPT and Planck cross-spectra over the SPT-SZ survey region. In this section, we turn our attention to how these simulations are created and used. The final product of this procedure is 400 sets of simulated unbiased bandpowers for each of the three combinations of data (150 $\times$ 150, 150 $\times$ 143, and 143 $\times$ 143).

3.2.1. Sky Signals

Our simulations include the following components: (1) a gravitationally lensed CMB signal, (2) thermal and kinematic Sunyaev–Zel'dovich (SZ) signals, (3) the cosmic infrared background (CIB) signal, and (4) emission from radio galaxies. We generate 400 realizations of the lensed CMB with LensPix (Lewis 2005), based on the best-fit ΛCDM model from Planck 2015 TT, TE, EE + lowP + lensing (Planck Collaboration et al. 2016b). Modes are generated out to $\ell_{\max} = 8000$, well above the angular multipoles, $\ell < 2500$, used in this comparison.

The maps are stored in HEALPix (Górski et al. 2005) format with resolution parameter $N_{\text{side}} = 8192$.

Unlike S13, which used Gaussian realizations for all extragalactic foregrounds, here we generate realizations of individual sources for the bright CIB galaxies (6.4 mJy $< S < 50$ mJy) and all radio galaxies (up to the flux
The bright CIB galaxies are drawn from the modeled $dn/dS$ of Cai et al. (2013), while the radio galaxy $dn/dS$ is taken from De Zotti et al. (2005). In both cases, the amplitudes of the $dn/dS$ distribution are calibrated by the actual observations of Mocanu et al. (2013) at 150 GHz.

The other extragalactic foregrounds, the thermal and kinematic SZ signals and the low-flux CIB, are treated as Gaussian. The shapes of the thermal and kinematic SZ angular power spectra are taken from the Shaw et al. (2010) and Shaw et al. (2012) models, respectively, with the amplitudes set to the median values from George et al. (2015), $D_{500}^{SZ} = 4.38 \mu K^2$ and $D_{300}^{SZ} = 1.57 \mu K^2$. We similarly draw upon the median values and templates from George et al. (2015) for the CIB terms. The clustered CIB spectrum is taken to follow $D_{\ell} \propto \ell^{0.8}$ with an amplitude $D_{500}^{\text{CIB}} = 3.46 \mu K^2$. The shot-noise or “Poisson” CIB power from galaxies dimmer than 6.4 mJy is taken to be $D_{300}^{\text{CIB}} = 9.16 \mu K^2$.

Each component is scaled appropriately to the effective frequency of the SPT or Planck maps and then co-added together to create the final sky realization.

### 3.2.2. Planck Noise Simulations

The Planck Collaboration released the 8th Full Focal Plane (FFP8) simulation in the 2015 data release (Planck Collaboration et al. 2016d). There are 1000 full-mission FFP8 noise simulations available at NERSC. We use the FFP8 full-mission simulations to create noisy Planck-like realizations to characterize the noise contribution to the SPT-Planck cross ($150 \times 143$) bandpowers as discussed in Section 4.2.

The FFP8 noise simulations are only available for the full-mission data and thus cannot be used for the Planck half-mission cross-spectrum. To create unbiased Planck 143 GHz bandpowers within the SPT area, we also generate half-mission noise simulations to characterize the noise contribution to the Planck–Planck ($143 \times 143$) bandpowers as presented in Section 4.2. The half-mission noise simulations are based on the pixel-space noise variance reported by the Planck Collaboration along with the half-mission observation maps. We approximate the Planck map noise as Gaussian and uncorrelated between pixels (“white”) for the half-mission noise simulations.

To test this white-noise approximation, we compare the noise power spectrum over the SPT survey area from 100 of these Gaussian white-noise realizations to 100 FFP8 full-mission noise simulations. The average power ratio is plotted in Figure 2. Clearly, the white-noise realizations overestimate the noise power for $150 < \ell < 2200$, i.e., most angular multipoles of interest. This implies that the Planck noise contribution is overestimated in our Planck-only bandpowers, at a maximum level of roughly 15% and a mean level of $\lesssim 10\%$ across the $\ell$ range of interest. We discuss the possible impact of this overestimate on our results in Section 4.2 and find that our main conclusions would likely be unchanged had we used more realistic noise simulations for the Planck-only bandpowers.

SPT noise realizations are created directly from the data. For each individual SPT observation of each SPT-SZ survey field, we create a residual map that is the difference between a map made from all left-going telescope scans and a map made from all right-going telescope scans. There are approximately 200 jackknife maps for each field (the number of individual $\sim 2$ hr observations). We use these maps to create the noise part of the simulated SPT observations. For each field, we multiply $+1$ or $-1$ randomly to the residual maps and co-add them together to form one noise realization. Using this method, we create 400 noise realizations.

With the same method, we also create SPT noise realizations of the first and second half sets of observations by co-adding the residual maps of half of the observations within one field. Similar to the half-mission noise simulations for Planck, we use the SPT noise simulations to characterize the noise contribution to the SPT-only bandpowers.

### 3.3. Filtering Transfer Function

The noise-free simulated observations are used to calculate $F_b$, the filtering transfer function, and $H_b$, the two-dimensional weight in the power spectrum calculation. Given the input power spectrum of the simulations, the effective transfer function can be derived by comparing the power spectrum of the simulations with the known input spectrum using an iterative scheme (Hivon et al. 2002). This method was used in K11 and S13; in this analysis, we make adjustments to Equations (6) and (7) in Keisler et al. (2011) to include the Planck beam and pixel window function for the SPT-Planck cross-spectrum and Planck-only spectrum.

### 3.4. Consistency Metrics

We have three sets of unbiased bandpowers derived from SPT and Planck data, $D_{b}^{50 \times 150}$, $D_{b}^{50 \times 143}$, and $D_{b}^{143 \times 143}$. In this analysis, we use the difference between these bandpowers, in particular the residual and the ratio between the bandpower sets, to quantitatively characterize the consistency between the SPT and Planck data sets. This section presents the metrics we use for these consistency tests.

---

59 https://crd.lbl.gov/departments/computational-science/c3/c3-research/cosmic-microwave-background/cmb-data-at-nersc/
3.4.1. Covariance Estimation

We estimate the bandpower sample variance and noise variance using 400 sets of simulated signal+noise maps. The final bandpower covariance matrix will also include a contribution due to beam uncertainties; the calculation of the beam covariance is detailed in the next section. For a single field and data combination, the bandpower covariance, $\Xi^{X \times Y}$, is calculated simply as

$$\Xi^{X \times Y} = \langle (D_{b,\text{sim}}^{X \times Y} - \bar{D}_{b,\text{sim}}^{X \times Y})(D_{b',\text{sim}}^{X \times Y} - \bar{D}_{b',\text{sim}}^{X \times Y}) \rangle,$$

where $\bar{D}_{b,\text{sim}}^{X \times Y}$ is the mean over all 400 simulations for cross-bandpowers $X \times Y \in \{150 \times 150, 150 \times 143, 143 \times 143\}$. The simulated $150 \times 143$ cross-bandpowers $D_{b,\text{sim}}^{150 \times 143}$ are derived from cross-spectra of simulated SPT maps including realizations of SPT full-observation noise and simulated Planck maps of the same underlying sky signal including Planck FFP8 noise simulations. For $\bar{D}_{b,\text{sim}}^{150 \times 150}$, the simulated bandpowers are calculated by the cross-correlation of the two sets of simulated SPT maps with half-observation noise realizations. Similarly, the simulated $D_{b,\text{sim}}^{143 \times 143}$ bandpowers are obtained using two sets of simulated Planck maps including Planck half-mission white-noise realizations. The full bandpower covariance matrix is then obtained by combining the individual-field covariance matrices with the square of the area weighting used to combine the bandpowers themselves.

As we will evaluate consistency by looking at the differences and ratios among sets of bandpowers, we also need the covariance matrix for these quantities. For the differences or residual bandpowers, $\Delta D_b = D_{b,\text{sim}}^{150 \times 150} - D_{b,\text{sim}}^{150 \times 143}$ (where $i, j, m, n \in \{150, 143\}$), the noise-plus-sample-variance part of the covariance is easily estimated from the 400 simulations:

$$\Xi_{\text{resid}} = \langle \Delta D_{b,\text{sim}}^{150 \times 150} \Delta D_{b',\text{sim}}^{150 \times 143} \rangle.$$

For the ratios relative to the $150 \times 150$ bandpowers, the covariance can be expressed as

$$\Xi_{\text{ratio}} = \left\langle \left( \frac{D_{b,\text{sim}}^{ij}^{150 \times 150}}{D_{b,\text{sim}}^{150 \times 150}} - 1 \right) \left( \frac{D_{b',\text{sim}}^{ij}^{150 \times 150}}{D_{b',\text{sim}}^{150 \times 150}} - 1 \right) \rightangle.$$

3.4.2. Beam Uncertainty

In this section, we present how beam uncertainties are handled for the bandpower comparison. Beam uncertainties appear as a second covariance term, in addition to the noise-plus-sample-variance term presented above, because the simulations did not include beam uncertainties.

We begin by estimating the eigenmodes of the Planck and SPT beam covariance matrices. The Planck beam eigenmodes can be obtained from RIMO in the 2015 data release, and the SPT beam eigenmodes can be derived from the beam correlation matrix calculated in S13. For each eigenmode, we can calculate the fractional beam uncertainty as a function of multipole, $\delta b_l / b_l$, and propagate this linearly to the bandpower space according to

$$\delta D_{b,\text{beam}}^{i \times j} = - \sum_{\ell} W_{b \ell}^{i \times j} D_{b \ell}^{i \times j, \text{fid}} \left( \frac{\delta b_j}{b_j} + \frac{\delta b_i}{b_i} \right),$$

where $i, j \in \{150, 143\}$, $W_{b \ell}^{i \times j}$ is the bandpower window function, and $D_{b \ell}^{i \times j, \text{fid}}$ is the fiducial power spectrum including CMB and extragalactic foregrounds for that frequency combination.

For the consistency tests, we need the beam covariance for bandpower differences and bandpower ratios. For the bandpower differences, the above equation leads straightforwardly to

$$\Xi_{bb, \text{beam}}^{(i \times j)(m \times n)} = \sum_{\text{e-modes}} (\delta D_{b,\text{beam}}^{150 \times 150} - \delta D_{b,\text{beam}}^{150 \times 143}) \times (\delta D_{b',\text{beam}}^{150 \times 150} - \delta D_{b',\text{beam}}^{150 \times 143}),$$

where $i, j, m, n \in \{150, 143\}$ and the sum is taken over all eigenmodes of the beam covariance matrices.

For the bandpower ratios, the beam covariance can be written as

$$\Xi_{bb, \text{beam}}^{(i \times j)(m \times n)} = \sum_{\text{e-modes}} \left( \frac{\delta D_{b,\text{beam}}^{150 \times 150} - \delta D_{b,\text{beam}}^{150 \times 143}}{D_i^{150 \times 150}} \right) \times \left( \frac{\delta D_{b',\text{beam}}^{150 \times 150} - \delta D_{b',\text{beam}}^{150 \times 143}}{D_i^{150 \times 143}} \right).$$

The total bandpower covariance for a consistency test is then the sum of the noise-plus-sample-variance term from the previous section and the appropriate beam covariance term above.

3.4.3. Bandpower Window Function Correction

To characterize the level of consistency between the three sets of bandpowers ($150 \times 150$, $150 \times 143$, and $143 \times 143$), we choose one set, $D_i^{150 \times 150}$, as the fiducial that is subtracted from or divided into the other two sets (see Section 4.2). However, the residual bandpowers and bandpower ratios are biased owing to the difference of bandpower window functions (e.g., Knox 1999). Due to the different filtering and beam transfer functions of the two experiments, the bandpower window functions are different between bandpower sets. In Figure 3, the window functions of the three sets of bandpowers are illustrated for two bins with effective centers at $\ell = 675$ and $\ell = 2475$. In both of these bins, relative to the $150 \times 150$ window functions, the $143 \times 143$ and $150 \times 143$ bandpowers
receive more weight from multipoles lower than the bin center and less from multipoles higher than the bin center. These differences, if left unaddressed, lead to a bias in the bandpower differences that is dependent, to some degree, on the assumed cosmological model.

Before comparing the bandpower sets, we need to correct for this bias. The correction is calculated as follows:

$$\delta D_b^i = \sum_\ell (W_\ell^i - W_\ell^{150} \tilde{D}_\ell^\text{fid}) D_\ell^i,$$

where $i = 150 \times 143$, or $143 \times 143$. The Planck 2015 best-fit cosmology plus the best-fit extragalactic foregrounds are used as the fiducial model. We subtract this correction term from the $150 \times 143$ and $143 \times 143$ bandpowers to remove the bias caused by the bandpower window function differences. In Figure 4, the top panel shows the shape of the correction $\delta D_b$ and the bottom panel shows the ratio between $\delta D_b$ and the corresponding $D_b$. For $150 \times 143$, $\delta D_b$ yields a roughly flat 1%–2% correction to the bandpower, while for $143 \times 143$, the correction is more important at higher multipoles, up to $\sim 15\%$ at $\ell = 2500$. The window function correction is clearly critical for the consistency analysis between these three sets of bandpowers.

We now investigate the model dependence of the window function corrections. To characterize this dependence, we calculate the window function corrections for a distribution of input $D_\ell$. The distribution is obtained by randomly sampling the cosmology and foreground models from a Markov chain based on WMAP9+Planck data. The variation of $\delta D_b$ in Equation (10) is illustrated in the top panel of Figure 4, with the error bars indicating the $3\sigma$ variation of the corrections. Within the band range of interest, this variation level is less than 1% of the bandpower standard error; therefore, we ignore this very small uncertainty in our analysis.

4. Results

In this section, we present the results of the comparison between SPT 150 GHz data and Planck 143 GHz data. We first present a qualitative map-level comparison of the two data sets in a common sky area. We then present the major result of the paper: the bandpower comparison among three sets of bandpowers calculated on the $2540 \text{deg}^2$ SPT-SZ sky patch: $D_b^{150\times150}$, $D_b^{150\times143}$, and $D_b^{143\times143}$. Setting $D_b^{150\times150}$ as the fiducial and subtracting it from the other two, we use the residual as the quantity to characterize the level of consistency between SPT 150 GHz and Planck 143 GHz data. As a by-product, for the first time we recalibrate the SPT data to the Planck data on the same patch of sky.

We also characterize the ratios between these sets of bandpowers. Note that the residual and ratio tests are not intended to be independent checks, as they contain nearly the same information presented in different ways. The reason we use both metrics is that the systematic errors most likely to cause differences between SPT and Planck data (such as unmodeled foreground residuals or beam systematics) fall into two broad categories: additive or multiplicative systematics. Additive systematics will show up most obviously in the residual bandpower test, while multiplicative systematics will show up most obviously in a bandpower ratio test.

Finally, we compare each set of $2540 \text{deg}^2$ bandpowers to the full-sky Planck 143 GHz bandpowers. Obviously such bandpower comparisons are not distinct from the map comparison. Rather, the residuals and the ratios of these sets of bandpowers are the quantitative version of the by-eye map comparison.

4.1. Map-level Comparison

The SPT and Planck maps within the same patch are presented in Figure 5 for visual comparison. The top panels show the filtered SPT map described in Section 2.1 on the left and the projected Planck full-mission 143 GHz map on the right. Some bright point sources can be identified by eye in both the SPT and Planck maps, but the maps do not resemble each other, because the Planck map is dominated by the degree-scale CMB anisotropy filtered out of the SPT data, and the SPT map shows more small-scale structure owing to its higher angular resolution. In the bottom panels of Figure 5, we show SPT and Planck maps of the same spatial modes. The bottom left panel shows the SPT map from the top left panel convolved with the difference between the SPT 150 GHz and Planck 143 GHz beams; the bottom right panel shows a map made by observing a Planck sky with the SPT 150 GHz scan strategy and TOD filtering. Though the Planck map has a visibly higher noise level (as expected), the signals in the two maps appear nearly identical. Figure 6 shows the difference between the bottom left and bottom right maps from Figure 5, on the same color scale, along with a simulated difference map.
for comparison. The feature at the location of one of the bright point sources is most likely caused by temporal variability of the source, as the SPT and Planck data were not taken simultaneously.

To make this result more quantitative, we could match the beam and filtering between these two data sets over the full SPT-SZ survey region, mask point sources, calculate the power in the difference map, and compare that power to the expected power from noise alone. In the next section we do a nearly equivalent but somewhat simpler calculation. Using the cross-spectrum formalism outlined in Section 3, we calculate the Planck 143 GHz power spectrum, the SPT 150 GHz power spectrum, and the SPT-Planck 150 × 143 cross-spectrum, and we calculate the χ² of the null hypothesis that these three spectra are measuring the same power. This calculation does not completely eliminate sample variance, as the perfectly mode-matched difference-map power spectrum would, but it strongly reduces it (see Figure 8 below).

### 4.2 Bandpower Comparison and Recalibration

In this section, we present the comparison between the three sets of window-function-corrected bandpowers, which we denote as \( D_b^{150 \times 150} \), \( D_b^{150 \times 143} \), and \( D_b^{143 \times 143} \). We first show all three sets of bandpowers in the top panel of Figure 7. The error bars contain contributions from sample variance, noise variance, and beam uncertainties. By eye, the three sets of bandpowers look very consistent. At low \( \ell \), the scatter among the three sets of points is much smaller than the errors on any one set; this is because the errors on all three are dominated by sample variance in this regime, and sample variance is highly correlated among the three data sets. This is also why the three sets of error bars are nearly the same size at low \( \ell \) (though the

---

**Figure 5.** Temperature maps of a portion of one SPT-SZ survey field. Top left: SPT-SZ 150 GHz map of this region. Modes with \( \ell \leq 600 \) are strongly suppressed in this map by the high-pass filter applied to the time-ordered data. Top right: Planck 143 GHz full-mission map of the same area. Bottom left: SPT map from the top left panel smoothed to have the same resolution as the Planck map. Bottom right: Planck map from the top right panel with the SPT-SZ high-pass filter applied. The difference between the bottom left and bottom right panels is shown in Figure 6. Note: the grayscale range of the top right panel is \([-300 \mu K, 300 \mu K]\), and the grayscale range of the other three panels is \([-100 \mu K, 100 \mu K]\).
143 × 143 bandpowers have slightly smaller error bars in this ℓ range because the modes lost in the SPT TOD filtering process result in slightly increased sample variance). Noise variance begins to dominate the 143 × 143 error bars at ℓ ≥ 1700 and the 150 × 143 error bars at ℓ ≥ 2000, while the 150 × 150 error bars are sample variance dominated over the entire range plotted.

The remaining panels of Figure 7 show various comparisons among the three sets of bandpowers. Two comparison schemes are applied: differences and ratios. The former is good for diagnosing additive effects, while the latter is more sensitive to multiplicative systematics. In both schemes, \( \mathbf{D}_b^{150 \times 150} \) is chosen as the fiducial bandpower set. These comparison plots and the \( \chi^2 \) and the probability to exceed (PTE) statistics calculated below are thus testing the following set of null hypotheses: (1) two sets of bandpower residuals

\[
\langle \Delta \mathbf{D}_{b,c} \rangle \equiv \langle \mathbf{D}_b^{150 \times 143} - \mathbf{D}_b^{150 \times 150} \rangle = 0
\]

\[
\langle \Delta \mathbf{D}_{b,a} \rangle \equiv \langle \mathbf{D}_b^{143 \times 143} - \mathbf{D}_b^{50 \times 150} \rangle = 0
\]

and (2) two sets of bandpower ratios

\[
\langle \Delta \mathbf{D}_{b,c}/\mathbf{D}_b^{150 \times 150} \rangle \equiv \langle \mathbf{D}_b^{150 \times 143}/\mathbf{D}_b^{150 \times 150} \rangle - 1 = 0
\]

\[
\langle \Delta \mathbf{D}_{b,a}/\mathbf{D}_b^{150 \times 150} \rangle \equiv \langle \mathbf{D}_b^{143 \times 143}/\mathbf{D}_b^{150 \times 150} \rangle - 1 = 0.
\]

The absolute calibration of the SPT data from George et al. (2015) has a statistical uncertainty of ∼1% at 150 GHz. We expect the bandpower comparison to be significantly more precise than this, so before plotting the bandpower residuals and ratios and before testing the null hypotheses above, we apply a recalibration parameter \( r_c \) to bandpowers containing SPT data:

\[
\Delta \mathbf{D}_{b,c} \equiv r_c \mathbf{D}_b^{150 \times 143} - r_c^2 \mathbf{D}_b^{150 \times 150}
\]

\[
\Delta \mathbf{D}_{b,a} \equiv \mathbf{D}_b^{143 \times 143} - r_c^2 \mathbf{D}_b^{50 \times 150}.
\]

and similarly for the bandpower ratios. Because of the precision of the George et al. (2015) calibration, we expect \( r_c \) to be very close to 1. We then calculate and minimize \( \chi^2 \) as a function of \( r_c \):

\[
\chi^2(r_c) = \Delta \mathbf{D}_b^T (r_c) \Xi^{-1} \Delta \mathbf{D}_b(r_c),
\]

and similarly for the bandpower ratios, where \( \Delta \mathbf{D}_b \) denotes the vector that includes both \( \Delta \mathbf{D}_{b,c} \) and \( \Delta \mathbf{D}_{b,a} \). We fit the two sets of residuals or ratios simultaneously. In principle, we should include the recalibration parameter in an adjustment to the noise contribution to the covariance matrix, but we neglect it for simplicity, with the justification that the correction is very small at less than 1%. In Figure 7, we have included the best-fit recalibration parameter for SPT, \( r_c \), in all three panels.

In the middle panel of Figure 7 we show the bandpower residuals \( \Delta \mathbf{D}_{b,c} \) and \( \Delta \mathbf{D}_{b,a} \) with error bars given by the square root of the diagonal elements of the full covariance matrix. This figure shows that the residual bandpowers are consistent with zero given the errors. In the bottom panels of the same figure we show the bandpower ratios. Similar to the residuals, the error bars of the bandpower ratios are the square root of the diagonal elements of the full covariance matrix. Qualitatively, these results appear to be consistent with our null hypotheses.

The quantitative characterization comes from the \( \chi^2 \) value with the best-fit recalibration parameter. The best-fit values of \( r_c \), minimum \( \chi^2 \), and probabilities to exceed that \( \chi^2 \) are listed in Table 1 for several combinations of data and covariance. The primary results are from the combined residual bandpowers. Using the full 2 × 2 block covariance matrix for the recalibration fit, these results give a best fit of \( r_c = 1.0087 \pm 0.0015 \) with \( \chi^2 = 78.7 \). There are 37 ℓ bins in our analysis, so we have 74 data points among the two residual bandpowers and one free parameter. The PTE for \( \chi^2 = 78.7 \) and 73 degrees of freedom is 0.30. We find very similar results...
from the combined-ratio fit: $r_c = 1.0092 \pm 0.0015$, $\chi^2 = 81.02$, PTE = 0.24. Put another way, given the noise properties and the beam uncertainties of the two experiments, 30% (24%) of our simulations have a higher $c^2$ for the bandpower differences (ratios) than we find with the real data. We thus find as our primary result that the SPT and Planck data in the SPT-SZ sky patch are quite consistent with the null hypothesis that there is no systematic offset in the two experiments’ measurement of the sky. A by-product of this analysis is that we recalibrate the SPT data with a statistical precision of 0.30% in power (0.15% in temperature) relative to Planck 143 GHz. This is comparable with the absolute calibration uncertainty of the Planck 143 GHz, 0.14% in power (0.07% in temperature; Planck Collaboration et al. 2016a), so we add this uncertainty in quadrature for a final SPT 150 GHz calibration uncertainty of 0.33% in power.

In Table 1, we also report the quantities of consistency from the single pair of bandpowers. For example, the PTE is 0.22 for the residual between $D_b^{150\times143}$ and $D_b^{150\times150}$ with the full covariance matrix. For the other pair of residual bandpowers, $D_b^{143\times143}$ and $D_b^{150\times150}$, the PTE is 0.43 with the full covariance matrix. The results without the beam uncertainties are also listed in Table 1.

As discussed in Section 3.2.2, the white-noise assumption in simulations for $D_b^{143\times143}$ results in an overestimate of the noise contribution to the covariance matrix, at roughly the 10% level in $D_b$ errors, or the 20% level in variance. If we assumed that the Planck noise was the dominant contribution to the residual and ratio covariance in the $150 \times 150$ versus $143 \times 143$ comparisons, we would expect roughly a 20% increase in $c^2$ if we were able to use more realistic noise simulations. The resulting $\chi^2$ would still correspond to a reasonable PTE ($\gtrsim 0.16$) for the null hypothesis. This would also be true of the combined constraints, particularly because the main constraining power comes from $D_b^{150\times143}$, which is unaffected by the assumption of white noise in the Planck-only bandpowers.
Because of the largely reduced sample variance contribution to the difference between bandpowers within the same sky coverage (roughly a factor of six in 150 × 143 − 150 × 150 compared to 150 × 150), the comparison of $D_b^{150×143}$ with the $D_b^{150×143}$ and $D_b^{143×143}$ bandpowers derived from within the SPT patch provides tighter tests, over a wide range of angular scales, than can be achieved by the comparison of $D_b^{150×150}$ with the more precise Planck spectra derived from nearly the full sky. In Figure 8, we compare the uncertainties on the two sets of bandpower residuals $D_b^{143×143} − D_b^{150×150}$ (blue) and $D_b^{150×143} − D_b^{150×150}$ (green) to the uncertainties in the SPT-only bandpowers $D_b^{150×150}$ (red). In the lower-ℓ region (ℓ < 1800), the green curve has the lowest error because the sample variance has been greatly reduced by subtracting the SPT bandpowers from the cross-bandpowers, while at higher ℓ the green curve rises owing to the Planck noise contribution. In the most constraining case ($D_b^{150×143} − D_b^{150×150}$), the errors for almost all bins are ≤10 μK$^2$, comparable to the uncertainty in the PlanckFS bandpowers. Both sets of bandpower residuals yield very stringent tests on the consistency of the two data sets, and our results show that these data sets are formally consistent in this patch.

### 4.3. Comparison to the Full-sky Planck 2015 TT High-ℓ Bandpowers

While restricting the bandpower comparison to overlapping sky substantially reduces differential sample variance, it does increase the Planck covariance. In this section, we instead investigate differences between the “in-patch” bandpowers (i.e., the 150 × 150, 150 × 143, and 143 × 143 bandpowers from the 2540 deg$^2$ SPT-SZ survey region) and the full-sky Planck 2015 TT high-ℓ unbinned frequency combined bandpowers (Planck Collaboration et al. 2016b), which we refer to as “PlanckFS.” In this case, the null hypothesis is that any differences between PlanckFS and the in-patch bandpowers are consistent with expectations given Gaussian statistics and statistical isotropy.

In Figure 9 we show the residuals and ratios for the in-patch to PlanckFS bandpowers. The residual plots show that all three sets of the in-patch bandpowers prefer greater power at 650 < ℓ < 1200, indicating that the SPT-SZ patch has greater power at these multipoles than the full-sky average. The ratio plots also suggest a tilt with respect to PlanckFS, although the significance is modest. The slope of the tilt is consistent to 0.5σ between the three cross-spectra, being slightly larger in 150 × 150 and smaller in 143 × 143.

To quantify the statistical significance of this tilt, we model the ratio of the in-patch and PlanckFS bandpowers as a power law:

$$\frac{D_b^{\text{in} \times j}}{D_b^{\text{PlanckFS}}} = A \left( \frac{\ell}{4000} \right)^n.$$  (14)

We assume a Gaussian likelihood with

$$-2 \ln \mathcal{L}(A, n) D_b^{\text{in} \times j} D_b^{\text{PlanckFS}} = \Delta_b X_{\ell=0}^{-1} \Delta_b',$$  (15)

$$\Delta_b = D_b^{\text{in} \times j} W_{bf}^{\text{in} \times j} D_b^{\text{PlanckFS}} - R_b^{\text{in} \times j}(A, n),$$  (16)

$$R_b^{\text{in} \times j}(A, n) = W_b^{\text{in} \times j} \left( A \left( \frac{\ell}{4000} \right)^n + F_{\ell}^{\text{in} \times j} \right).$$  (17)

Here $D_b^{\text{PlanckFS}}$ is the best-fit PlanckFS power spectrum, $\Sigma_{bb'} = \langle \Delta_b \Delta_{b'} \rangle$ (only the SPT beam uncertainty is included and cross-correlations between $D_b^{\text{in} \times j}$ and the $D_b^{\text{PlanckFS}}$ are negligible), and $F_{\ell}^{\text{in} \times j}$ is the foreground model adopted from S13 with the frequency dependence from George et al. (2015) included. There are three foreground amplitude parameters included in $F_{\ell}^{\text{in} \times j}$ to account for the SZ, Poisson-point-source, and clustered CIB uncertainties. The best-fit recalibration from Section 4.3 has been applied to the in-patch SPT-only and SPT-Planck bandpowers. The best-fit values for $A$ and $n$ are reported in Figure 9. We find only marginal evidence for a tilt between PlanckFS and any of the three cross-spectra; the most significant tilt is for 150 × 150 and is 1.5σ away from zero. We conclude that the tilts we see in the observed spectra are roughly consistent with the expectation based on Gaussian
statistics and the assumption of statistical isotropy—i.e., they are roughly consistent with \( n = 0 \).

To determine the expected statistical properties of the differences in the best-fit values of \( A \) and \( n \), we construct a covariance matrix \( \mathbf{C} \) from the set of best-fit values of \( A \) and \( n \) in each of the 400 simulations of Section 3.2. We then calculate \( \chi^2 \) for these differences as 

\[
\chi^2 = (\mathbf{\Theta} - \bar{\mathbf{\Theta}})^T \mathbf{C}^{-1} (\mathbf{\Theta} - \bar{\mathbf{\Theta}}),
\]

where \( \mathbf{\Theta} \) is the vector of parameter differences and \( \bar{\mathbf{\Theta}} \) is the mean simulation difference, which is consistent with zero. The breakdown of PTEs for the various pair differences is shown in Table 2. As can be seen, the most extreme PTE is 0.85 and the lowest is 0.41. We conclude that the observed tilts in the three cross-spectra are completely consistent with each other.

All our tests are consistent with the following explanation for the tilts we observe in the in-patch spectra relative to the best-fit \( \text{Planck} \) full-sky spectrum: they are driven by a sample variance fluctuation away from the full-sky average, the magnitude of which is roughly consistent with expectations under the assumption of statistical isotropy and our noise model.

### 4.4. Pipeline Checks

All published power spectra from the SPT Collaboration have been calculated using some variant of the cross-spectrum pseudo-\( C_\ell \) pipeline used in this work. Extensive checks have been performed on this pipeline (see, e.g., Section 4.2 of Story et al. 2013), demonstrating that the correct input spectrum is recovered from simulated data, even when that spectrum differs from the spectrum assumed in calculating the filter transfer function. If, however, some aspect of the pipeline were inducing a bias in the estimated power spectra (through some mechanism that has escaped all pipeline tests), this bias would affect both the SPT-SZ and \( \text{Planck} \) data used in this work (because we have mock-observed the \( \text{Planck} \) data and analyzed it with the SPT pipeline). If the bias on the two data sets were comparable, it would then divide or subtract out in the bandpower comparison, and we would (wrongly) conclude that there was no issue with either data set.

To test this scenario, we have analyzed the in-patch \( \text{Planck} \) data using an alternate pipeline. Specifically, we created a HEALPix version of the SPT-SZ sky-patch and point-source mask (stitched together from the individual-field masks), and we handed this mask and the half-mission \( \text{Planck} \) 143 GHz maps to the PolSpice$^{40}$ (Szapudi et al. 2001; Chon et al. 2004) code, which is designed to estimate the power spectra of masked full-sky maps and properly account for the masking. We binned the \( \Delta \ell = 1 \) PolSpice output into the \( \Delta \ell = 50 \) bins used in the SPT pipeline using the bandpower window functions calculated for the \( 143 \times 143 \) “scanned, filtered” \( \text{Planck} \) bandpowers. We then calculated the ratio of the “unscanned, unfiltered” (PolSpice) bandpowers to the scanned, filtered ones and found that they agree to better than 3% in every individual bin in which there is appreciable signal-to-noise ratio in the \( 143 \times 143 \) spectrum, with an overall ratio of \( 1.0028 \pm 0.0050 \) over the range \( 600 < \ell < 1800 \) and no evidence of a trend with \( \ell \). We have also redone the tilt calculation in Section 4.3 using the unscanned, unfiltered bandpowers and found results consistent with what we found with the scanned, filtered bandpowers (within a fraction of a sigma). We thus conclude that our fundamental results are not an artifact of the SPT analysis method.

### 5. Conclusions

In this paper, we have compared 150 GHz SPT data and 143 GHz \( \text{Planck} \) data in the same region of the sky, namely, the 2540 deg$^2$ SPT-SZ survey footprint. We have performed a visual comparison of maps constructed from the two data sets and found the difference between the two maps to be visually consistent with noise, once they have been filtered to display the same angular modes.

We then performed a quantitative analysis of the consistency of the maps, relying primarily on a comparison of the

---

$^{40}$ http://www2.iap.fr/users/hivon/software/PolSpice
cross-spectrum of two halves of the SPT data with the SPT × Planck cross-spectrum. We also compared the SPT × SPT spectrum with the cross-spectrum of two halves of the Planck data. These comparisons were made using differences between and ratios of two sets of binned power spectra, or bandpowers, at a time, always using the SPT × SPT bandpowers as the fiducial set. To test the null hypothesis that the bandpower differences (after recalibrating the SPT data) are consistent with zero—or that the ratios are consistent with unity—we created a suite of 400 simulations of the signal and noise properties of the SPT and Planck maps, including signal contributions from the CMB and extragalactic foregrounds. We found our most stringent test, based on the expected variance of the differences, to be the comparison of the SPT × SPT spectrum with the SPT-Planck cross-spectrum. Forming a \( \chi^2 \) quantity from these bandpower differences and the simulation-based covariance matrix, we have found a value that is exceeded by 22\% of the analogous \( \chi^2 \) values derived from the simulated data, i.e., corresponding to the PTE value 0.22. When we add the residuals between the SPT × SPT and Planck × Planck cross-spectra, we find a PTE of 30\%. All other tests result in similarly unremarkable PTEs. We find no evidence of a failure of our null model, i.e., we see no evidence for systematic errors or under- or overestimates of statistical errors.

We have also compared the three sets of bandpowers from the 2540 deg\(^2\) SPT-SZ survey region to the full-sky Planck 143 GHz power spectrum. Relative to the Planck full-sky spectrum, we have found a hint for a tilt in the in-patch bandpowers. For all three sets of in-patch bandpowers, the amplitudes of the tilts we have obtained are consistent with each other and roughly consistent with expected noise and sample variance fluctuations.

This work shows that the SPT 150 GHz and Planck 143 GHz data are in very good agreement with each other within the 2540 deg\(^2\) SPT-SZ survey area. In a companion paper (Aylor et al. 2017), we extend this comparison to the cosmological parameters that can be derived from these bandpowers.

The South Pole Telescope is supported by the National Science Foundation through grant PLR-1248097. Partial support is also provided by the NSF Physics Frontier Center grant PHY-1125897 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation, and the Gordon and Betty Moore Foundation grant GBMF 947. The McGill group acknowledges funding from the National Sciences and Engineering Research Council of Canada, the Canada Research Chairs program, and the Canadian Institute for Advanced Research. Argonne National Laboratory work was supported under U.S. Department of Energy contract DE-AC02-06CH11357. B.B. is supported by the Fermi Research Alliance, LLC under contract No. De-AC02-07CH11359 with the U.S. Department of Energy. C.R. acknowledges support from an Australian Research Council Future Fellowship (FT150100074). We thank D. Rapetti for comments on the manuscript, and we thank an anonymous referee for helpful suggestions.

**Appendix**

In this appendix, we explain why we use a different power spectrum estimator for the SPT bandpowers than in S13, and we discuss the potential difference this makes to our comparisons. S13 averaged the cross-spectra between \( O(100) \) single-observation maps, while in this work we use the cross-spectrum between two half-survey maps in each field. Similarly, S13 estimated the bandpower covariance matrix from the distribution of said cross-spectra, while we estimate the covariance from signal+noise simulations.

The decision to change estimators is driven by the desire to use the same procedure for the 150 × 150, 150 × 143, and 143 × 143 bandpowers. With only the two half-survey maps for 143 GHz, the S13 covariance estimator would not work. Instead, we rely on signal+noise simulations to calculate the bandpower covariance. However, we could only create one noise realization per map for the single-observation maps used in S13. By using half-survey maps (i.e., by co-adding many single-observation maps), we can increase the number of independent noise realizations dramatically. As laid out in Section 3.2.3, we generate noise realizations by first nulling the signal in single-observation maps by differencing the left-going and right-going scans and then co-adding these noise maps with a random +1 or −1 prefactor. These noise realizations can then be added to the simulated signal-only maps to yield robust signal+noise SPT map simulations.

One might worry that the change in estimator could affect our comparison. To address this, we perform the following quantitative test. We replace the SPT 150 GHz half1-half2 cross bandpower by the original S13 bandpowers and redo the test on the bandpower residuals (with the covariance matrix of the residual unchanged). We find no significant differences. For \( \mathcal{D}_{150}^{143-150-150} \), with only the S13 bandpowers, the best-fit \( r_c = 1.0090 \pm 0.0016 \) with \( \chi^2 = 47.21 \) and PTE = 0.100. Recall from Table 1 that the numbers with the half-survey bandpowers are \( r_c = 1.0087 \pm 0.0016 \) with \( \chi^2 = 42.07 \) and PTE = 0.224. For \( \mathcal{D}_{150}^{143-143-150} \), with the S13 bandpowers, the best-fit \( r_c = 1.0086 \pm 0.0022 \) with \( \chi^2 = 39.17 \) and PTE = 0.330. Again from Table 1, the equivalent numbers with the half-survey bandpowers are \( r_c = 1.0076 \pm 0.0022 \) with \( \chi^2 = 36.83 \) and PTE = 0.430. We therefore conclude that the change in SPT power spectrum estimators does not significantly impact our comparison of the SPT 150 GHz and Planck 143 GHz map over the same region of sky.

**ORCID iDs**

T. M. Crawford @ https://orcid.org/0000-0001-9000-5013
T. de Haan @ https://orcid.org/0000-0001-5105-9473
E. M. George @ https://orcid.org/0000-0001-7874-0445
G. P. Holder @ https://orcid.org/0000-0002-0463-6394
D. P. Marrone @ https://orcid.org/0000-0002-2367-1080
Y. Omori @ https://orcid.org/0000-0002-0963-7310
C. L. Reichardt @ https://orcid.org/0000-0003-2226-9169
A. A. Stark @ https://orcid.org/0000-0002-2718-9996
J. D. Vieira @ https://orcid.org/0000-0001-7192-3871

**References**

Addison, G. E., Huang, Y., Watts, D. J., et al. 2016, ApJ, 818, 132
Aylor, K., Hou, Z., Knox, K. T., et al. 2017, ApJ, 850, 101
Battye, R. A., & Moss, A. 2014, PhRvL, 112, 051303
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
Cai, Z.-Y., Lapi, A., Xia, J.-Q., et al. 2013, ApJ, 768, 21
Calabrese, E., Hlozek, R. A., Battaglia, N., et al. 2013, PhRvD, 87, 103012
Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, MNRAS, 350, 914
