

THE ATACAMA COSMOLOGY TELESCOPE: MEASUREMENT AND ANALYSIS OF 1D BEAMS FOR DR4

MARIUS LUNGU¹, EMILIE R. STORER¹, MATTHEW HASSELFIELD², ADRIAAN J. DUIVENVOORDEN¹,
ERMINIA CALABRESE³, GRACE E. CHESMORE⁴, STEVE K. CHOI^{5,6}, JO DUNKLEY^{1,7}, ROLANDO DÜNNER⁸,
PATRICIO A. GALLARDO⁶, JOSEPH E. GOLEC⁴, YILUN GUAN⁹, J. COLIN HILL^{2,10}, ADAM D. HINCKS¹¹,
JOHANNES HUBMAYR¹², MATHEW S. MADHAVACHERIL^{13,14}, MAYA MALLABY-KAY¹⁵, JEFF MCMAHON^{4,15,16,17},
KAVILAN MOODLEY^{18,19}, SIGURD NAESS²⁰, FEDERICO NATI²¹, MICHAEL D. NIEMACK^{5,6,22}, LYMAN A. PAGE¹,
BRUCE PARTRIDGE²³, ROBERTO PUDDU⁸, ALESSANDRO SCHILLACI²⁴, CRISTÓBAL SIFÓN²⁵, SUZANNE STAGGS¹,
DHANESHWAR D. SUNDER^{18,19}, EDWARD J. WOLLACK²⁶, ZHILEI XU^{27,28}

Draft version February 21, 2022

ABSTRACT

We describe the measurement and treatment of the telescope beams for the Atacama Cosmology Telescope’s fourth data release, DR4. Observations of Uranus are used to measure the central portion ($< 12'$) of the beams to roughly -40 dB of the peak. Such planet maps in intensity are used to construct azimuthally averaged beam profiles, which are fit with a physically motivated model before being transformed into Fourier space. We investigate and quantify a number of percent-level corrections to the beams, all of which are important for precision cosmology. Uranus maps in polarization are used to measure the temperature-to-polarization leakage in the main part of the beams, which is $\lesssim 1\%$ (2.5%) at 150 GHz (98 GHz). The beams also have polarized sidelobes, which are measured with observations of Saturn and deprojected from the ACT time-ordered data. Notable changes relative to past ACT beam analyses include an improved subtraction of the atmospheric effects from Uranus calibration maps, incorporation of a scattering term in the beam profile model, and refinements to the beam model uncertainties and the main temperature-to-polarization leakage terms in the ACT power spectrum analysis.

¹ Joseph Henry Laboratories of Physics, Jadwin Hall, Princeton University, Princeton, NJ 08544, USA

² Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010 USA

³ School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, Wales CF24 3AA, UK

⁴ Department of Physics, University of Chicago, Chicago, IL 60637, USA

⁵ Department of Physics, Cornell University, Ithaca, NY 14853, USA

⁶ Department of Astronomy, Cornell University, Ithaca, NY 14853, USA

⁷ Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA

⁸ Instituto de Astrofísica and Centro de Astro-Ingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile

⁹ Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George St., Toronto, ON M5S 3H4, Canada

¹⁰ Department of Physics, Columbia University, New York, New York 10027, USA

¹¹ David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St. George St., Toronto ON M5S 3H4, Canada

¹² NIST, Quantum Sensors Group, 325 Broadway, Boulder, CO 80305, USA

¹³ Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo ON N2L 2Y5, Canada

¹⁴ Department of Physics and Astronomy, University of Southern California, Los Angeles, CA, 90007, USA

¹⁵ Department of Astronomy and Astrophysics, University of Chicago, 5640 S. Ellis Ave., Chicago, IL 60637, USA

¹⁶ Kavli Institute for Cosmological Physics, University of Chicago, 5640 S. Ellis Ave., Chicago, IL 60637, USA

¹⁷ Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA

¹⁸ Astrophysics Research Centre, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa

¹⁹ School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa

²⁰ Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029, Blindern, NO-0315 Oslo, Norway

²¹ Department of Physics, University of Milano - Bicocca, Piazza della Scienza, 3 - 20126, Milano (MI), Italy

²² Kavli Institute at Cornell for Nanoscale Science, Cornell University, Ithaca, NY 14853, USA

²³ Department of Physics and Astronomy, Haverford College, Haverford, PA 19041, USA

²⁴ Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA

²⁵ Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile

²⁶ NASA/Goddard Space Flight Center, Greenbelt, MD 20771, USA

²⁷ Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

²⁸ MIT Kavli Institute, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

1. INTRODUCTION

The Atacama Cosmology Telescope (ACT) is a 6 m off-axis Gregorian telescope located at an altitude of 5190 m in the Atacama Desert of northern Chile. It is designed for millimeter wavelength observations of the cosmic microwave background (CMB) at arcminute resolution. The telescope and receiver are described in Fowler et al. (2007) and Thornton et al. (2016) respectively. This paper describes the treatment of the ACT beams for its fourth data release, referred to as DR4. This data release includes temperature and polarization data collected by ACT between 2013 and 2016, covering seven regions of the sky spanning roughly $18,000 \text{ deg}^2$, in frequency bands centered on 98 and 150 GHz (Thornton et al. 2016). The DR4 data were obtained using two 150 GHz detector arrays, PA1 and PA2, and one dichroic detector array, PA3, operating at 98 and 150 GHz. The three array positions are not identical optically. Each year of data is referred to as a season (S13–S16) and was analyzed separately. The DR4 data products and some of their analyses are presented in Choi et al. (2020), Aiola et al. (2020), Darwish et al. (2020), Han et al. (2021), Madhavacheril et al. (2020), Namikawa et al. (2020), and Mallaby-Kay et al. (2021).

The beams of the telescope determine the instrument’s response to different scales on the sky. Quantifying these beams and the uncertainties on the beam measurements is one of the most challenging aspects of characterizing the instrument. An incorrect calibration of the beams would bias virtually all of the ACT science analyses, including the precision measurement of the CMB power spectrum.

The previous paper that focused on ACT beams, Hasselfield et al. (2013), dates from DR2, which included data through 2010. The DR3 analysis papers (Louis et al. 2017; Naess et al. 2014) referenced Hasselfield et al. (2013) and described the small changes in methodology since then. In recent years, there have been several modifications to our beam pipeline that merit further discussion. This paper provides a comprehensive guide to the analysis of the ACT beams for DR4.

The paper is organized as follows. In §2 we describe the planet observations used to measure the main portion of the ACT beams. In §3 we explain how these observations are made into maps. In §4 we describe the steps of the beam pipeline from planet maps to a model of the ACT beams and their covariance. §5 accounts for the polarization component of the beams. Finally, in §6 we discuss the publicly released beam products, assumptions made in the analysis, and future directions for ACT beam analyses. Part of the information in this paper overlaps with that in previous ACT papers, including Aiola et al. (2020), Choi et al. (2020), and Hasselfield et al. (2013), and is included here for clarity and completeness.

2. OBSERVATIONS

Observations of sufficiently bright point-like objects are required in order to properly characterize the shape of the beams out to large angles. For ACT, planets are the best candidates for this purpose, though not all equally so. The proximity of Mercury and Venus to the Sun make them difficult to observe at night; Mars’ temperature is not sufficiently constant (see, for example, Wright

1976), so it intermittently becomes too bright; Jupiter’s apparent size is too large, compared to the instrument’s angular resolution, to be considered a point source (and it is too bright, as is Saturn); Saturn, though suitable for sidelobe analysis, is bright enough to induce a non-linear detector response near peak amplitude, resulting in signal suppression within the main lobe of the beam;¹ Neptune’s small angular diameter relative to the beamwidth dilutes its brightness, reducing signal-to-noise. We have thus chosen to base the majority of the ACT beam analysis on observations of Uranus, which achieves adequate signal-to-noise without exceeding the dynamic range of the instrument and can be approximately treated as a point source.² As described in §5.2, Saturn is used to study the beam’s polarized sidelobes.

Planet observations are done by briefly interrupting the CMB survey scans, changing the azimuth at which the telescope is pointing, and scanning again at a fixed elevation. Uranus is observed roughly once per night during the observing season, but only the higher quality observations are used to characterize the beam. Before making maps of the observations, we apply the same data selection criteria as for the main maps used for cosmological analysis. For example, we discard data taken during the daytime, data taken when the weather was particularly bad, or data from detectors that do not meet criteria summarized in Aiola et al. (2020) and detailed in Dünner et al. (2013). Once maps of the observations are made, further selection criteria are applied, as described in §4.1.

3. PLANET MAP-MAKING

3.1. Filtering And Noise

Uranus maps were made using a special filter-and-bin map-maker built on the same framework used for building our normal maximum-likelihood maps (`enki`).³ The reason we do not use maximum-likelihood maps for planets is that these are not robust to “model errors”, as described in e.g. Naess (2019). The difference between the true continuous sky and the pixelated model is interpreted as noise, leading to a bias which propagates along the scan direction for a noise correlation length. This is a significant effect for a bright, localized signal such as a planet.

It is possible to make maximum-likelihood models robust to these effects at significant computational and implementational cost (Piazzo 2017), but we here choose a simpler method that exploits the compactness of the planet signal relative to the atmospheric noise correlation length to separate the two. The idea is to split the data into two regions: a target region (including the planet) which we want to map, and a reference region (the rest

¹ Such detector non-linearity may have contributed to the $\sim 7\%$ discrepancy between the *Planck* and ACT DR2 measurements of Saturn’s temperature (Planck Collaboration Int. LII et al. 2017; Hasselfield et al. 2013).

² While not implemented for the current beam analysis, another possibility being studied involves using observations of Uranus to measure the central portion of the beam (where Saturn would induce a non-linear response) and observations of Saturn to measure the parts of the beam further from the peak. This way, Saturn’s brightness allows us to measure the outer parts of the beam to greater signal-to-noise than when using observations of Uranus alone.

³ GitHub repository: <https://github.com/amaurea/enlib>

of the data) which is used to estimate the behavior of the atmospheric noise.⁴ The atmosphere in the target region is then estimated by interpolating the atmospheric modes from the reference region, and it is then subtracted from the target region of the map.⁵ To the extent that the signal in the target region is uncorrelated with any signal in the reference region, then this subtraction will not introduce bias on any scale in the target region. A point source like Uranus is very close to fulfilling these criteria, but in practice the outermost parts of the beam extend into the reference region, introducing a slight bias that we address in §3.2.

The target region must be large enough to capture all of the main beam, but small enough that the atmosphere’s behavior in the target region can be adequately predicted using only data from the reference region. We choose a 12’ radius area centered on the planet as the target region, and estimate the atmosphere’s behavior here by exploiting its spatial correlation structure. Nearby detectors see similar parts of the atmosphere, so if at one moment in time some detectors are seeing the target region and some are seeing the reference region, the data from the latter are used to help determine the atmospheric modes in the former.⁶ This new interpolation technique significantly improves the noise modelling compared to the technique used in [Louis et al. \(2017\)](#), which used a simpler variant of filter-and-bin where the detector common mode was subtracted.

The resulting planet maps are cleaner and dominated mostly by white noise. In addition, we now inverse variance weight the detectors when making the planet maps, which was not done previously. This improved mapping has contributed to a more detailed understanding of the temperature-to-polarization leakage, which will be described in §5.1.

3.2. Bias From Map-Making

While it is approximately true that all the planet signal is contained inside the target region, a small part of it extends into the reference region due to the faint wings of the beams. This introduces a small but non-negligible bias in the target region when the atmosphere is subtracted.

To study this bias, a set of simulated beam-convolved Uranus observations are processed through the map-maker. The beam-convolved planet signal is simulated using a 2D beam model⁷ which uses the inverse Fourier transforms of two-dimensional Zernike polynomials as its basis set, similar to the 1D model of the beam described in §4.2.3. This simulated planet signal is then injected into several subsets of real time-domain data taken from CMB observations. The resulting data are

⁴ The atmospheric fluctuations seen by ACT are described in more detail in [Morris et al. \(2021\)](#).

⁵ For a previous implementation of a map-making method designed to model and subtract the atmosphere from maps of planets, see [Hincks et al. \(2010\)](#).

⁶ Ideally one would also use the temporal correlations of the atmosphere, and solve the most probable value of the atmosphere in the target region given the data in the reference region. This technique may be employed in the future. In preliminary tests, this only results in a mild improvement.

⁷ This model was developed as part of a larger, ongoing project to fit the beams in two dimensions in order to capture their asymmetry.

then processed through the planet map-making pipeline, which includes estimating and subtracting the atmospheric noise. Both the CMB observations alone and the CMB observations with the planet signal injected are mapped separately, and then the former are subtracted from the latter. The result is a set of noise-free simulated maps of Uranus observations.⁸

An example of a map of a simulated Uranus observation is shown in Figure 1. Radial profiles are constructed by taking the azimuthal average of the input beam model and the mean output of the simulated observations. The difference between these two radial profiles represents the bias due to the atmospheric subtraction. As seen in Figure 2, other than small residual fluctuations near the center of the beam, the bias can be well approximated by a constant offset.⁹ In these simulations the target region for atmosphere subtraction was chosen to have a 16’ radius, but the qualitative conclusions we draw are applicable to the final planet maps used for the beam analysis which have a target region with a 12’ radius.¹⁰ As shown in Figure 2, the effect of the atmosphere removal is almost exactly as expected: the value of the input beam at the target radius (16’ in this case) is uniformly subtracted within that radius to produce the output beam.

So the planet map-making bias is well approximated by a constant offset in each map, determined by the value of the Uranus signal at the edge of the target radius used for noise estimation. This offset is estimated and subtracted prior to analysis, as described in §4.2.1.

4. BEAM PIPELINE

4.1. Map Selection And Processing

A map of Uranus is made for each detector array, frequency, and observation.¹¹ Only a fraction of these maps are used in the final beam analysis. We select maps made from observations done at night (23–11 UTC) taken with the secondary mirror in its final focus position for the season in question. We pick the maps where there are not too many zero-weight pixels (which are indicative of poor scan coverage), where the signal-to-noise is sufficiently high (> 10), and where there are not too many horizontal stripes (determined by visual inspection of the maps). This striping in the maps is caused by large-scale variations due to residual atmosphere. The rejection of maps with too much striping is new, and is done to simplify the analysis. It results in fewer maps being selected

⁸ In the end, it was discovered that including the atmospheric noise in the simulations does not affect the results, since the atmosphere removal is approximately linear.

⁹ These small residual features only appear in the T maps when solving for the maps in T , Q , U simultaneously. When we solve only for T and set $Q = U = 0$, the fluctuations disappear. This seems to be due to an issue with the conditioning of the covariance matrix in the simulations, and is not a physical effect.

¹⁰ During preliminary beam studies, maps of both Uranus and Saturn were made, with target regions both of radius 12’ and 16’. Saturn was found to be too bright. In the Uranus maps with a target region of radius 16’, the signal-to-noise in the radial profiles past 12’ was insufficient to be of use in the analysis. Hence, the Uranus maps with a target region of radius 12’ were used, in order to make the best possible prediction of the atmosphere in the region of interest.

¹¹ So we do not make any per-detector maps. For each detector array, frequency, and observation of Uranus, we combine the data from all the detectors that meet the quality criteria mentioned in §2 into one map.

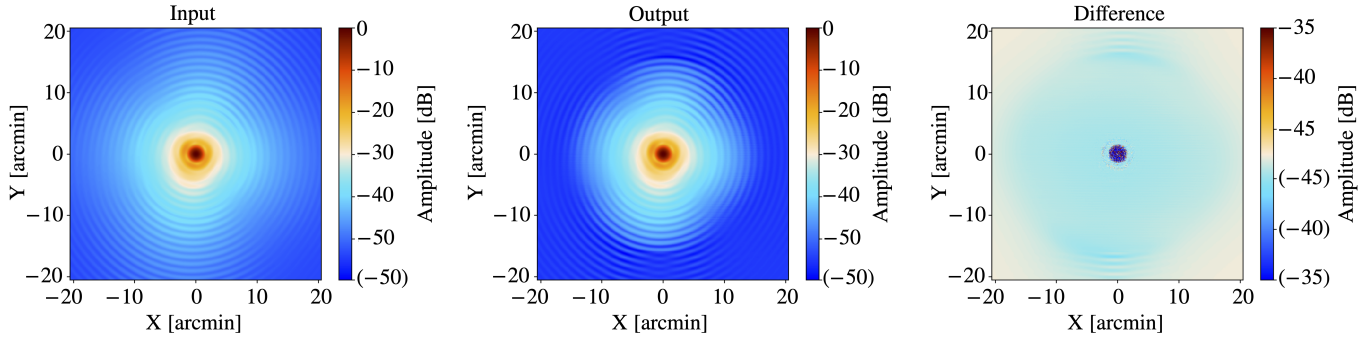


FIG. 1.— Example of a simulated planet observation used to characterize the mapping transfer function. This example is for S16 PA2 at 150 GHz. (Left) The input 2D beam model. (Center) One simulated observation. (Right) The difference between the “input” beam model (left) and the map of the “output” simulated observation (center). Note that in each case the color scale is a symmetrical log scale^a in dB (with a linear threshold of -50 dB for the input and output maps and -45 dB for the difference map) and negative values are enclosed by parentheses. Here the target region where the atmosphere was estimated and removed during the map-making process was chosen to have a radius of $16'$. The large-scale residuals seen are mostly constant within this target region of the map. The slight asymmetry in the beam (which may be safely ignored for DR4, as explained in §6.2) is due primarily to the positions of the telescope’s optics tubes in the focal plane.

^aA symmetrical log scale is logarithmic in both the positive and negative directions, with a small linear range around zero to avoid having values tend to infinity (https://matplotlib.org/stable/api/scale_api.html#matplotlib.scale.SymmetricalLogScale).

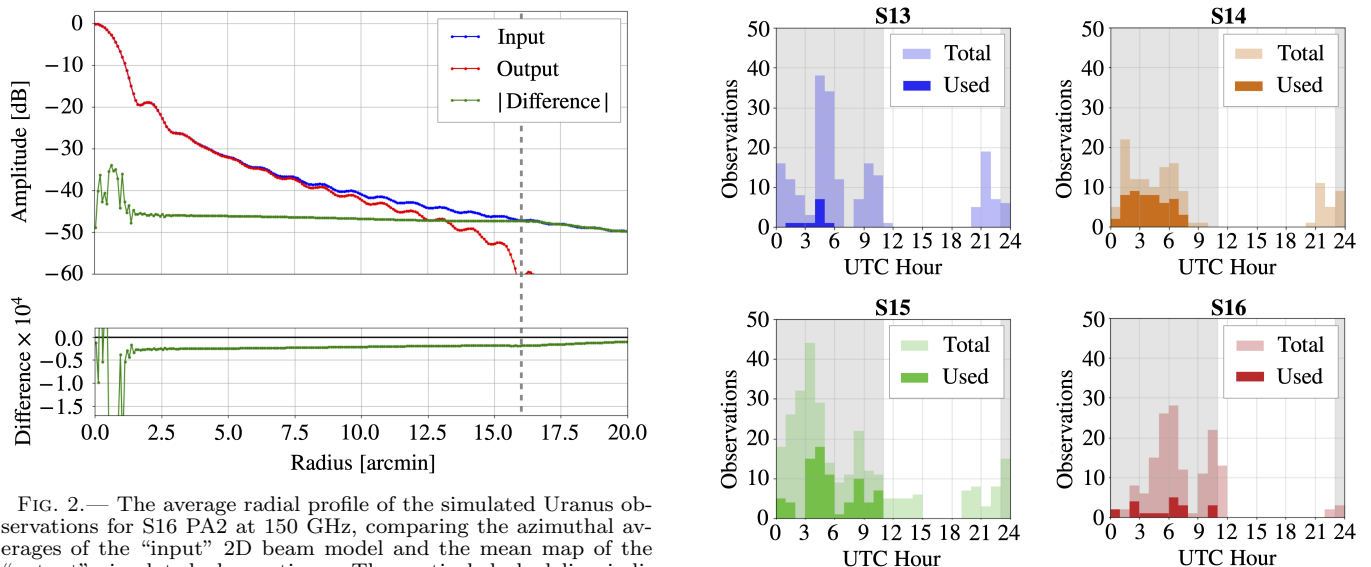


FIG. 2.— The average radial profile of the simulated Uranus observations for S16 PA2 at 150 GHz, comparing the azimuthal averages of the “input” 2D beam model and the mean map of the “output” simulated observations. The vertical dashed line indicates the $16'$ radius target region of the map within which the atmosphere was estimated and subtracted. (For the Uranus maps used to characterize the beam, a $12'$ radius target region is used, but the qualitative conclusions we draw here still apply.) Other than the small (roughly -40 to -30 dB) variations near the beam center, the difference is well approximated by a constant offset in the region from $3.5'$ to $10.0'$ where we fit for it. As described in §4.6, we adjust our beam covariance matrices to account for possible uncertainty due to variations in the range over which we fit the offsets, exploring the three independent ranges of $3.5'$ – $5.0'$, $5.0'$ – $7.0'$, and $7.0'$ – $10.0'$.

than in Louis et al. (2017). For each season, array, and frequency, the number of maps discarded due to striping ranges from 3 to 42 (on average 19), which represents between 4% and 36% of the maps (on average 16%).

The number of Uranus maps used for the beam analysis versus the total number of observations is shown in Figure 3 and Table 1. For example, in 2014 (S14) we made 129 observations of Uranus. In the case of PA2, we discarded 34 observations because at the time the telescope pointing was optimized for PA1 (so Uranus was not well

FIG. 3.— Distribution of the total number of Uranus observations that were made (faded colors) and ultimately became part of the final beam analysis (solid colors) for all arrays combined, shown by observing seasons from 2013–16. The shaded regions between 0 UTC and 11 UTC, as well as 23 and 24 UTC, demarcate the ACT nighttime dataset (note that local time at the observing site fluctuates between UTC-3 and UTC-4). See Table 1 for a summary of the number of observations used vs total for each detector array and season.

measured by PA2), 36 observations because the resulting maps had too much striping, and 8 because the signal-to-noise was too low. This leaves 51 Uranus observations to measure the beam. An example “good” map which was selected for the beam analysis and a “bad” map which was discarded due to too much striping are shown in Figure 4. In S13, a large fraction of the Uranus maps were thrown out because the observations were made in the early commissioning phase of the telescope, before it had achieved its final focus.

For the purposes of the beam measurements, we re-center each Uranus map by fitting a 2D Gaussian in the

TABLE 1
SUMMARY OF URANUS OBSERVATIONS - NUMBER USED/TOTAL.

Array	Band (GHz)	Season	Used	Total
PA1	150	S13	11	197
		S14	45	129
		S15	17	133
PA2	150	S14	51	129
		S15	38	164
		S16	11	86
PA3	150	S15	8	117
		S16	6	78
PA3	98	S15	33	117
		S16	9	78

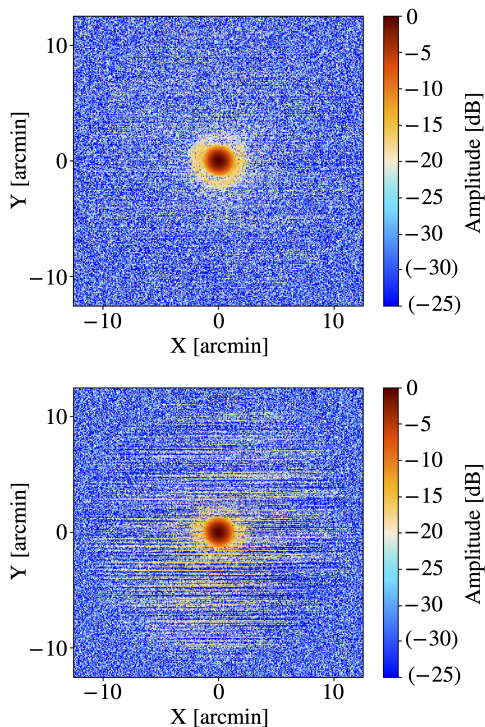


FIG. 4.— Two maps of individual Uranus observations for S15 PA2 at 150 GHz. (Top) A “good” map which was selected for the beam analysis. (Bottom) A “bad” map which was thrown out due to having too many horizontal stripes (caused by large-scale variations due to residual atmosphere remaining after the atmosphere subtraction described in §3.1). In each case, the color scale is a symmetrical log scale in dB, with a linear threshold of -30 dB and negative values enclosed by parentheses. Any slight asymmetry visible here may be safely ignored for DR4, as explained in §6.2, and is due primarily to the positions of the telescope’s optics tubes in the focal plane.

vicinity of the planet. The amplitudes from these fits are used to normalize the beam profiles to have peak values of unity. We also estimate the noise level in each map by computing the standard deviation of the pixel values outside the target region (the $12'$ radius area centered on the planet, described in §3.1).

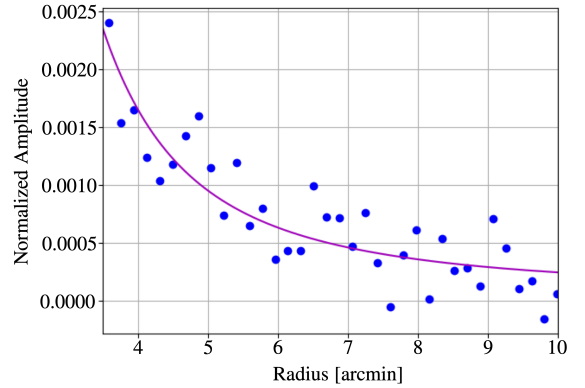


FIG. 5.— Example of the offset fit for the radial profile of one Uranus observation for S14 PA2 at 150 GHz. The points are found by taking the azimuthal average of the map for each radial bin. Uncertainties on these points are not computed. The curve is the best-fit model $\alpha/\theta^3 + \beta$, where in this case $\alpha = 0.08934$ and $\beta = 0.00012$.

4.2. Radial Profile Fitting

4.2.1. Radial Profiles

The instantaneous ACT beams are slightly elliptical. However, when fitting the beam we work with azimuthally averaged (“symmetrized”) radial beam profiles, treating the beams as if they were circular. In our case, the cross-linking of the scans (visible in Figure 2 of Choi et al. 2020) means that the telescope beams contribute to the maps with roughly equal weight at two different orientations that are approximately 90 degrees apart for a large part of the ACT sky coverage. The effective beams are thus circularized.¹²

This symmetrizing effect works well in temperature, but it does not help with the temperature-to-polarization leakage caused by beam asymmetry. However, this effect has been quantified and may be safely ignored for DR4, as explained in §6.2.

The ACT beams are small enough that we use a flat sky approximation for modeling them. We denote each beam map by $B(\theta, \phi)$, where θ is the radial distance from the beam center and ϕ is the polar angle. We set $B(0, \phi) = 1$. The symmetrized radial beam profile is then

$$B(\theta) = \frac{1}{2\pi} \int d\phi' B(\theta, \phi'). \quad (1)$$

Each map, with a target region of radius $12'$, is binned into a symmetrized radial profile with bins of width $11''$ out to a radius of $10'$, for a total of 55 bins. The dominant source of noise in the Uranus maps comes from variations in the atmosphere which occur at relatively large angular scales, so there is significant correlation between the bins in each radial profile.

Before combining the radial profiles into one mean profile for each season, array, and frequency, each map’s profile must be corrected for the offset due to the bias

¹² In reality, anisotropic noise weighting (due to ACT’s noise being correlated along the scan direction) makes the beam symmetrization from the map-maker is different from what one gets by simply averaging. The overall effect is to either make the circularized beam smaller or larger than the naive prediction, depending on the direction the raw beam is elongated relative to the scanning direction. This effect is absorbed into the jitter correction described in §4.5.

induced by the planet map-maker that was described in §3.2. This is done by fitting the model $\alpha/\theta^3 + \beta$ (motivated in §4.2.4) plus a scattering term (described in §4.2.5) to each profile in the range 3.5'–10.0' and then subtracting the measured offset, β , from each. An example of such a fit is shown in Figure 5.

We then compute an average radial profile independently for each season, detector array, and frequency. When averaging radial profiles, each profile is weighted by $1/N$ where N is the white noise variance estimated from the corresponding map. The averaged radial profiles extend to roughly -40 dB of the peak, or 35 dBi (40 dBi) at 150 GHz (98 GHz).

An advantage of computing each radial profile then taking the average, compared to making one average Uranus map then computing its radial profile (as done in Hincks et al. (2010), for example), is that it more easily allows for the estimation of the full covariance matrix for the averaged radial profile. The off-diagonal elements of this matrix are important since there is significant noise correlation between the radial bins, which propagates into the low- ℓ uncertainty of the beam window function described in §4.3.

4.2.2. Radial Profile Covariance Matrix

As the averaged radial profile is computed for each season, array, and frequency, the covariance matrix between the amplitudes of the 55 radial bins is also computed in order to account for the covariant uncertainty on large angular scales. Given the small sample size, with only between 6 and 51 profile measurements used to estimate each matrix, a shrinking algorithm (Schäfer & Strimmer 2005) is used to down-weight the off-diagonal terms of the matrix. This method is described in detail in Appendix A. In our case this shrinking procedure is necessary as a standard estimate of the covariance matrix is often so noisy it becomes ill-conditioned and cannot be inverted.

The shrinkage technique works by combining an empirical estimate of the covariance matrix \mathbf{S} (a high-dimensional estimate of the underlying covariance with little or no bias) with a model \mathbf{T} (a low-dimensional estimate which may be biased but has much smaller variance) to minimize the total mean squared error (sum of bias squared and variance) with respect to the true underlying covariance:

$$\mathbf{C} = \hat{\lambda}^* \mathbf{T} + (1 - \hat{\lambda}^*) \mathbf{S}. \quad (2)$$

Here $\hat{\lambda}^*$ is the parameter (often referred to as shrinkage intensity) that determines the contribution of each matrix. In our case, the covariance matrix \mathbf{S} is an unbiased empirical estimate of the covariance, the sample covariance matrix, and the model matrix \mathbf{T} is given by the diagonal of \mathbf{S} :

$$T_{ij} = \begin{cases} S_{ii} & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases} \quad (3)$$

a common choice. We analytically calculate the optimal combination of the low and high dimensional estimates by determining $\hat{\lambda}^*$ from \mathbf{S} :

$$\hat{\lambda}^* = \frac{\sum_{i \neq j} \widehat{\text{Var}}(S_{ij})}{\sum_{i \neq j} S_{ij}^2}, \quad (4)$$

where $\widehat{\text{Var}}(S_{ij})$ is an estimate of the variance of each covariance matrix element. The derivation of $\hat{\lambda}^*$ can be found in Appendix A. For the analysis presented here, $\hat{\lambda}^*$ ranges from 0.24 to 1.

4.2.3. Core Model

Since ACT's primary mirror has a diameter of $D = 6$ m, at a given wavelength λ the diffraction limited optical response of the telescope is restricted to spatial frequencies below $\ell_{\text{max}} \simeq 2\pi D/\lambda$. The Fourier transform of each beam is therefore compact on a disk. Thus, a natural choice of basis functions to model the beams in harmonic space is the Zernike polynomials, an orthonormal set of basis functions that is complete on the unit disk.¹³

Expressed in polar coordinates ρ and ϕ in the aperture plane (where ρ is the radial distance $0 \leq \rho \leq 1$, and ϕ is the azimuthal angle $0 \leq \phi \leq 2\pi$), the Zernike polynomials may be written as

$$V_n^m(\rho, \phi) = R_n^m(\rho) e^{im\phi}, \quad (5)$$

where m and n are integers such that $n \geq |m|$ and $R_n^m(\rho)$ are a set of radial functions defined as

$$R_n^{\pm m}(\rho) = \begin{cases} \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! (\frac{n+m}{2} - k)! (\frac{n-m}{2} - k)!} \rho^{n-2k} & \text{for } n-|m| \\ & \text{even} \\ 0 & \text{for } n-|m| \text{ odd.} \end{cases} \quad (6)$$

For an azimuthally symmetric beam, it is only necessary to consider the polynomials for which $m = 0$, which may be expressed as

$$R_{2n}^0(\rho) = P_n(2\rho^2 - 1), \quad (7)$$

where $P_n(x)$ are Legendre polynomials (Born & Wolf (1980)). The Fourier transform of these polynomials is

$$\tilde{R}_{2n}^0(\theta) = \int_0^1 \rho d\rho e^{-i\rho\theta} R_{2n}^0(\rho) = \frac{(-1)^n J_{2n+1}(\theta)}{\theta}, \quad (8)$$

where J_{2n+1} is a Bessel function of the first kind.

Motivated by this, and following previous analyses (Hasselfield et al. 2013), we adopt the basis functions

$$f_n(\theta \ell_{\text{max}}) = \frac{J_{2n+1}(\theta \ell_{\text{max}})}{\theta \ell_{\text{max}}} \quad (9)$$

to fit the core of the beams in real space, where the parameter ℓ_{max} varies the scale of the function's argument and n is a non-negative integer.

4.2.4. Asymptotic Behavior

On scales smaller than a few arcminutes, the noise is subdominant to the planetary signal, and so the beam profiles are well measured even by a single observation of Uranus. However, at larger angular scales, the noise

¹³ The Zernike polynomials are usually used to fit effects in the electric field, not in intensity, but since they are a basis set, we can use them to fit the intensity beams. So the asymptotic behavior of f_n in Equation 9 is $1/\theta^{1.5}$, not $1/\theta^3$ as we expect for the beams (as described in §4.2.4).

quickly becomes non-negligible. Thus, the asymptotic behavior of the beams cannot be separated from the background without making some assumptions about the shape of the beams far from the core.

The illumination of the ACT optics is controlled by a cryogenic Lyot stop. The beam's Fraunhofer diffraction pattern for monochromatic radiation is described by the squared modulus of the Fourier transform of this circular aperture (Born & Wolf 1980). The resulting intensity response with angle, or Airy pattern, is

$$A(\theta) = \left[\frac{2J_1[k(D/2) \sin \theta]}{k(D/2) \sin \theta} \right]^2, \quad (10)$$

where $k = 2\pi/\lambda$ is the wavenumber.¹⁴ For ACT, $k(D/2) = \pi D/\lambda$ is large, since the primary is several thousand wavelengths across. When the argument of the Bessel function is $\gg 1$ and real, as is the case for $\theta > 2'$, we can make the approximation

$$J_1(x) = \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{3\pi}{4}\right) \quad (11)$$

(Abramowitz & Stegun 1972). In this regime, the Airy pattern asymptotically approaches

$$A(\theta) \simeq \frac{8 \cos^2[k(D/2) \sin \theta - 3\pi/4]}{\pi[k(D/2) \sin \theta]^3}, \quad (12)$$

such that the envelope of $A(\theta)$ falls as $1/\sin^3 \theta$. This implies that at angles larger than the ‘‘near sidelobes’’ (described in §5.2) and neglecting the effects of scattering, the beams should fall as $1/\sin^3 \theta \simeq 1/\theta^3$, since θ is small in the region considered for the beam models.

A different, more general, way to derive this asymptotic behavior of the beams comes from the geometric theory of diffraction from Keller (Keller 1956, 1959, 1962). As detailed in Page et al. (2003), for an illuminated disk, the diffraction pattern from the rim, regardless of the interior, leads to a gain (response) of the shape

$$G(\theta) \propto \frac{1}{\sin \theta} \left(\frac{1}{\sin \theta/2} \pm \frac{1}{\cos \theta/2} \right)^2 \quad (13a)$$

$$\propto \frac{1}{\sin^3 \theta} (1 \pm \sin \theta), \quad (13b)$$

for an electric field parallel (+) or perpendicular (−) to the edge, in the far field-limit. For unpolarized light, one can simply average over both polarizations.

The beam behavior derived here assumes perfectly smooth and uniform surfaces within the telescope (i.e. perfect coherence). Nonuniformity, including mirror surface roughness, leads to diffuse scattering (i.e. decoherence of the fields), and thus a reduction in the telescope gain (Ruze 1966). As described in the next section, an additional term is included in our beam model to account for this effect.

4.2.5. Scattering

The primary telescope surface can be characterized by an rms surface roughness, ϵ , and a correlation length,

¹⁴ Note that this corresponds to the f_0 term from the core model described in §4.2.3.

c . We measured these properties for ACT using photogrammetry,¹⁵ placing 284 targets on the corners of the primary's 71 panels where the panels join. The rms of the raw measurements was found to be approximately $30 \mu\text{m}$. However, these measurements are not a representative sample of the rms of the surface, since the targets are at the most extreme points on the corners of the panels. By fitting a polynomial surface to the measurements and looking at the residuals over the entire surface, we find an rms of $\epsilon = 20 \mu\text{m}$. Using the photogrammetry software, the uncertainty on the rms measurements is estimated to be $10 \mu\text{m}$. This does not include any misalignments due to macroscopic deformations of the telescope.

From these measurements, we can also compute the correlation length, c . The correlation function is

$$C(d) = \sum_{ij} z(\vec{r}_i) z(\vec{r}_j) = C_0 e^{-d^2/c^2}, \quad (14)$$

where \vec{r} is the position on the surface, z is the mean-subtracted departure from the best-fit designed surface (described in Section 3 of Fowler et al. 2007), $d = |\vec{r}_i - \vec{r}_j|$ is the separation, and the sum is over all measurement pairs with a separation d . We find that the shape of the correlation function follows the above form reasonably well. By fitting the photogrammetry measurements with Equation 14, for the ACT primary, we find $c = 280 \text{ mm}$. This scattering leads to another angular scale (or shoulder) in the beam response, not governed by the diameter of the Lyot stop.

For a beam $B(\theta)$ normalized to unity at $\theta = 0$, the corresponding solid angle Ω is

$$\Omega = 2\pi \int_0^\pi B(\theta) \sin \theta d\theta. \quad (15)$$

The expression for the gain due to scattering off a rough surface, $G(\theta)$, given by Equation 8 in Ruze (1966)¹⁶ can be re-written in terms of the unit-normalized beam for an ideal reflector, $B_0(\theta)$, its corresponding solid angle, Ω_0 , the total beam, $B(\theta)$, and its corresponding solid angle, Ω , by using the relation $G(\theta) = (4\pi/\Omega)B(\theta)$. We then obtain the equation

$$B(\theta) = \frac{\Omega}{\Omega_0} B_0(\theta) e^{-\langle \delta^2 \rangle} + S(\theta), \quad (16)$$

with the $S(\theta)$ term given by

$$S(\theta) = \frac{\Omega}{4\pi} \left(\frac{2\pi c}{\lambda} \right)^2 e^{-\langle \delta^2 \rangle} \sum_{n=1}^{n=\infty} \frac{\langle \delta^2 \rangle^n}{n \cdot n!} e^{-(c\pi \sin(\theta)/\lambda)^2/n}, \quad (17)$$

where $\langle \delta^2 \rangle = (4\pi\epsilon/\lambda)^2$ is the variance in the surface phase (and ϵ and c are the surface rms and correlation length, as described earlier).

The sum converges quickly, with four terms sufficient for our purposes. We have simulated the ACT beam by taking the Fourier transform of a perturbed surface with

¹⁵ VSTARS by GSI, <https://www.geodetic.com/v-stars/>

¹⁶ There is a typo in this equation in the original paper. Inside the summation, the variance δ^2 should be raised to the n^{th} power, as written here, instead of being simply squared, as in Ruze (1966) (Corkish 1990).

deformations of a Gaussian shape and verified that we indeed recover the Ruze equation as it is written above.¹⁷

The first term in Equation 16 above is expected to decay as $1/\theta^3$, but the same does not hold true for $S(\theta)$, so the inclusion of this term in our fitting procedure is important as it affects how the beam fits are extrapolated. This scattering term $S(\theta)$, which we refer to as the Ruze beam, contains roughly 1.5% of the solid angle of the main beams at 150 GHz. The addition of this term to the beam model is new to the DR4 analysis.

4.2.6. Beam Model Fitting

We separate the main beam for each season, array, and frequency into two domains: the core, where we fit the set of basis functions described in §4.2.3, and the wing, where the model is composed of a $1/\theta^3$ term in addition to the Ruze beam (where we use the measured values for ϵ and c described in the previous section). As explained in §4.2.1, a constant offset has already been subtracted from each measured beam profile before combining them into an average radial profile for each season, detector array, and frequency. The full model for each averaged beam profile can be written as

$$B(\theta) = \begin{cases} \sum_{n=0}^{n_{\text{mode}}-1} a_n f_n(\theta \ell_{\text{max}}) & \text{for } \theta \leq \theta_1 \\ \alpha/\theta^3 + S(\theta) & \text{for } \theta_1 < \theta. \end{cases} \quad (18)$$

The initial linear least squares fits are performed for the a_n and α for a range of values of ℓ_{max} , θ_1 , and n_{mode} (allowing n_{mode} to vary up to a maximum value of half the number of data points in the core). We first sample a range of values for the scaling parameter ℓ_{max} and then allow this non-linear parameter to vary later on. Based on the optics of the system, at 150 GHz (98 GHz) we expect to have $\ell_{\text{max}} \simeq 2\pi D/\lambda \simeq 19,000$ (12,500). We sample ℓ_{max} in integer steps from 1 to 30,000.

We fit the model out to $10'$, a value which is chosen because this is the radius out to which the beam is measured to sufficient signal-to-noise. Despite the atmosphere subtraction in the target region that extends out to $12'$, residual modes remain, so the signal-to-noise out to $12'$ is not sufficient. By only fitting out to $10'$ we ensure that the fit remains in the region where the measurement is dominated by the beam, not atmosphere.

The model is fit to each averaged radial profile, with its associated covariance matrix. We do not include continuity conditions at θ_1 , but any small discontinuities are well below ACT's resolution.¹⁸

¹⁷ More specifically, deformations of a Gaussian shape of width (standard deviation) 250 mm were placed on a square grid of dimension 580 mm on a side to create deformations over a surface. This gave 71 squares inside a diameter of 5500 mm, the effective diameter of ACT's primary. This is a reasonable approximation given that the primary has 71 panels (Thornton et al. 2016). The amplitudes of the deformations were drawn from a normal distribution, and the overall rms of the surface was adjusted to be $\epsilon = 100 \mu\text{m}$, and the modeled surface had $c = 440$ mm. The resulting beam followed the Ruze equation for these values of ϵ and c .

¹⁸ For example, the fit shown in the top panel of Figure 6 may have a slight discontinuity in amplitude at θ_1 . However, as part of the beam processing we take the harmonic transform of our beam model and then transform back to radial space to obtain the final

In order to select the optimal number n_{mode} for each beam profile, we must strike the right balance between minimizing the χ^2 and not adding too many parameters. To do this, we use the corrected Akaike information criterion (AICc), which estimates the relative quality of models based on both their goodness of fit and their simplicity. For each choice of ℓ_{max} , θ_1 and n_{mode} , the AICc is computed for the best-fitting model.

The uncorrected AIC is given by

$$\text{AIC} = 2k - 2 \ln(\hat{\mathcal{L}}), \quad (19)$$

where k is the number of estimated parameters in the model and $\hat{\mathcal{L}}$ is the maximum value of the likelihood function for the model (Akaike 1973, 1974). For small sample sizes, as is the case here, the AIC can exhibit a large bias. To account for this, we use the corrected criterion AICc, where

$$\text{AICc} = \text{AIC} + \frac{2k^2 + 2k}{n - k - 1}, \quad (20)$$

and n is the sample size (Hurvich & Tsai 1989).

We select the values for ℓ_{max} , θ_1 , and n_{mode} corresponding to the lowest AICc.^{19,20} For the DR4 beams, the best-fit values for ℓ_{max} range from 15,604 to 17,936 (from 10,960 to 11,285) at 150 GHz (98 GHz), the best-fit values for θ_1 range from $3.58'$ to $7.79'$, and the best-fit values for n_{mode} range from 9 to 13. We then use MCMC to sample the posterior distribution of the parameters ℓ_{max} , the a_n , and α , assuming uniform priors (Metropolis et al. 1953). This method produces an estimate of the parameter means and the full covariance between all the basis functions and wing parameters, including the non-linear scaling parameter ℓ_{max} .

The term $S(\theta)$ in Equation 18 depends on the beam's solid angle, Ω , but the calculation of Ω depends on the beam model. To get around this issue, the beam fit is performed iteratively, first using a rough estimate for Ω , and then re-computing Ω once a beam model is obtained, and then performing the fit again. In total, we fit the beam four times, re-estimating Ω each time. By the last iteration, the change in Ω becomes undetectable, and so we have converged to a final value for the beam solid angle.

We did test whether the data prefer an asymptotic wing fit term that differs from α/θ^3 , by fitting for α/θ^b instead. While the best-fitting value for b often differs (by 10–20%) from three, it is not strongly constrained by the data. The AICc indicates that the addition of this new parameter is not warranted.

An example of the beam model fit to radial profile data is shown in Figure 6, for the S15 PA2 Uranus maps at 150 GHz, together with the residuals. The core func-

beam profiles shown in Figure 8. Any small discontinuity that may have been present in the initial fit disappears due to the resolution with which we do the transform.

¹⁹ In reality, this is not always precisely the case. When there exists a set of values with n_{mode} smaller by one (hence one fewer parameter) and an AICc that is not significantly different ($\Delta\text{AICc} < 2$), we actually select the model with the slightly higher AICc. We do this to avoid over-fitting the data, since in this case the AICc does not justify the addition of the extra parameter.

²⁰ In earlier analyses (Louis et al. 2017; Hasselfield et al. 2013), we simply increased n_{mode} until the reduced- χ^2 fell below 1.

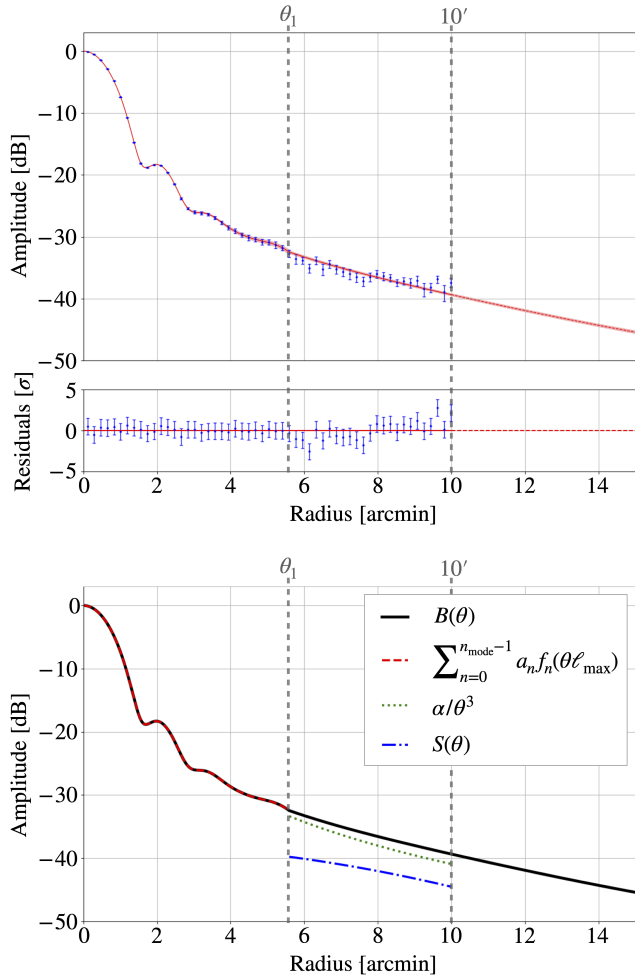


FIG. 6.— Top: Measured radial beam profile (blue) for S15 PA2 at 150 GHz and the best-fit model (red), with the (narrow) red shaded region indicating the 1σ model uncertainty bounds, and residuals shown in the middle panel. Note that the bins are correlated. Bottom: The terms in the model for the beam. Here the final best-fit value of θ_1 , where the model for the beam changes, is $5.59'$. The number of modes in the core fit (n_{mode}) is 13.

tions, α/θ^3 term, and the Ruze beam are indicated in the lower panel, with the core functions used out to a radius of $\theta_1 = 5.59'$, then transitioning to the α/θ^3 plus Ruze beam model at larger radii, and using this model to extrapolate past $10'$. We measure the beam profiles down to -40 dB from the peak, which would leave a few percent of the beam's solid angle unaccounted for if we did not use our fit to extrapolate past $10'$.

Figure 7 shows how the symmetrized model compares to the average Uranus map data for this case. The full set of radial profile fits for the S15 data is shown in Figure 8 (with some small additional corrections, as described in §4.4). We also obtained model fits for the S13, S14 and S16 data that make up DR4. The solid angles, gains, and FWHMs for all seasons are reported in Table 2 (again, with the small corrections from §4.4). For reference, $1 \text{ nsr} \simeq 0.0118 \text{ arcmin}^2$.

Even though we make a high signal-to-noise measurement of each beam, the uncertainty on the solid angle is limited to $\sim 2\%$ because of the uncertainty in extrapolation, which depends on the model. The fractional uncer-

TABLE 2
INSTANTANEOUS BEAM SOLID ANGLES, GAINS, AND FWHMS

Array	Band (GHz)	Season	Solid Angle (nsr)	Forward Gain (dBi)	FWHM (arcmin)
PA1	150	S13	201.5 ± 3.8	77.94 ± 0.08	1.330 ± 0.001
		S14	198.5 ± 3.3	78.01 ± 0.07	1.330 ± 0.002
		S15	196.5 ± 8.8	78.06 ± 0.19	1.321 ± 0.002
PA2	150	S14	183.1 ± 3.2	78.37 ± 0.08	1.310 ± 0.001
		S15	187.8 ± 4.7	78.26 ± 0.11	1.311 ± 0.001
		S16	185.4 ± 4.8	78.31 ± 0.11	1.316 ± 0.001
PA3	150	S15	269.6 ± 5.5	76.68 ± 0.09	1.461 ± 0.002
		S16	237.5 ± 8.5	77.24 ± 0.16	1.444 ± 0.003
PA3	98	S15	503.9 ± 21.8	73.97 ± 0.19	2.001 ± 0.004
		S16	476.3 ± 21.8	74.21 ± 0.20	2.002 ± 0.004

tainties in the solid angles at 98 GHz may be larger than those at 150 GHz in part because at 98 GHz the beam is broader, so a larger fraction of the beam is affected by the uncertainty in our extrapolation.

When using the beam solid angle to calculate the flux density of a point source, the uncertainty on the measurement may be reduced by first applying a matched filter to both the beam and the source in order to remove the scales associated with the extrapolation uncertainty.

While much of this fitting method follows the approach used in Hasselfield et al. (2013) and Louis et al. (2017), notable improvements are the addition of the scattering beam term in the model, the use of MCMC sampling that includes estimating the non-linear parameter ℓ_{max} , the exploration of different radii (θ_1) out to which the core model is fit, and the use of the AICc to choose the final best-fit model.

4.3. Beam Window Functions

In spherical harmonic space, the beam information is encoded in the harmonic transform b_ℓ and the window function $w_\ell = b_\ell^2$, which describes the instrument's response to different multipoles, ℓ . This window function is an essential component of the DR4 power spectrum analysis in Choi et al. (2020).

The harmonic transform b_ℓ is the Legendre transform, or more accurately the Legendre polynomial transform, of the beam radial profile:

$$b_\ell = \frac{2\pi}{\Omega} \int_{-1}^1 B(\theta) P_\ell(\cos \theta) d(\cos \theta). \quad (21)$$

For small beams such as that of ACT, this is effectively a Fourier transform. The derivation of the Legendre transform and details about how the transform is computed are presented in Appendix B.

We use b_ℓ instead of B_ℓ to indicate the division by Ω , which normalizes b_ℓ to unity at $\ell = 0$ (since $P_0 = 1$). $B_\ell = \Omega b_\ell$ has units of sr, whereas b_ℓ is dimensionless. We extrapolate the model beyond the fit radius of $10'$ when computing the transform. This is necessary to capture the low- ℓ part of the window function, and to account for the part of the beam solid angle that is beyond the range we fit.

A subset of the beam transforms from DR4 is shown in Figure 9. A similar figure in Aiola et al. (2020) shows window functions, b_ℓ^2 , which are used to correct the power

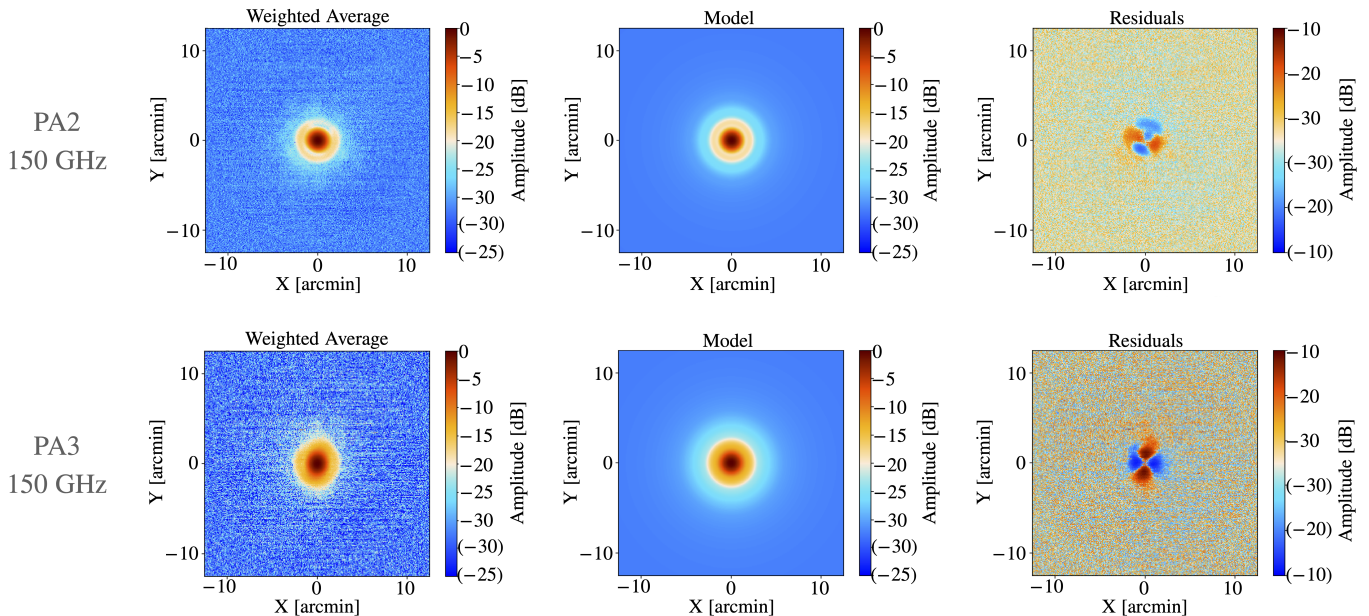


FIG. 7.— Maps of the main beam for S15 PA2 (top) and PA3 (bottom) at 150 GHz. In each case, the color scale is a symmetrical log scale in dB, with a linear threshold of -30 dB and negative values enclosed by parentheses. (Left) The weighted average of the Uranus observations used to characterize the beam. (Center) The model which is fit to the radial profile of the measured beam. (Right) The difference between the measured beam (left) and the beam model (center). Note that the color scale for the residuals is different from the other two maps. Both PA2 and PA3 are shown here, since PA2 serves as an example of a less elliptical beam, and PA3 serves as an example of a more elliptical one. The corresponding maps for S15 PA1 at 150 GHz resemble the ones shown for PA2 here, and the maps for S15 PA3 at 98 GHz are similar to the ones shown here for PA3 at 150 GHz, but broader. For each individual detector array, the residuals are fairly constant from one season to the next. As shown in Figure 6 for PA2, the azimuthal average of the residuals is consistent with zero, which is why the fit is successful, despite the residuals visible in the maps. These residuals are expected to have a quadrupole-like shape, since we know our beams are slightly elliptical, and the quadrupole is the dominant asymmetric azimuthal mode for an elliptical beam.

spectra. For a given array and season, if the beam transforms for PA1 and PA2 are, respectively, b_ℓ^{PA1} and b_ℓ^{PA2} , then for the auto-power spectrum of the PA1 or PA2 maps the window functions are $(b_\ell^{\text{PA1}})^2$ and $(b_\ell^{\text{PA2}})^2$, and for the cross-spectrum of the PA1 and PA2 maps the window function is $b_\ell^{\text{PA1}}b_\ell^{\text{PA2}}$.

4.4. Additional Corrections

The resulting beam models and covariance matrices are an accurate description of the binned radial beam profiles, but they must be corrected for some systematic effects. Following the same approach as in Hasselfield et al. (2013), corrections are applied to account for the pixelization of the planet maps, the binning of the maps into radial annuli, Uranus’ angular diameter, and the planet’s effective frequency. The effect of each of these corrections is shown in Figure 10 for S15 PA2 at 150 GHz, as a typical example.

To correct for the pixelization of the planet maps, we divide the beam transform by the azimuthal average of the pixel window function, $\text{sinc}(pk_x/2\pi)\text{sinc}(pk_y/2\pi)$, where $\{k_x, k_y\}$ are the spatial frequencies and p is the pixel size in radians. This is a $\lesssim 0.1\%$ effect for $\ell < 10,000$.

Then, we estimate and correct for the transfer function induced due to binning the planet maps into radial annuli. This is done by simulating planet maps, using the best-fitting beam profile model as the input, and then estimating their radial profiles following the same procedure as with the data. Comparing the input model with the output radial profile gives an estimate for the transfer function resulting from the radial binning. This is a

$\lesssim 1\%$ effect for $\ell < 10,000$.

We also correct for Uranus’ angular diameter, since Uranus is large enough that it cannot be treated as a point source given the precision to which we measure the beam. For each season, array, and frequency, we assume Uranus is a disk with radius equal to the weighted mean of the radii for all the Uranus observations contributing to the beam measurement. We then deconvolve this shape in harmonic space using the function $2J_1(lr)/lr$, where r is the radius of Uranus’ disk in radians. The factor of 2 normalizes the function to unity as $\ell \rightarrow 0$. This is a $\lesssim 0.1\%$ effect for $\ell < 10,000$.

The beam at this point describes the telescope’s response to a point-like source with an approximately Rayleigh-Jeans (RJ) spectrum. Near our frequencies, the temperature spectrum of Uranus goes roughly as $\nu^{-0.25}$ (Planck Collaboration Int. LII et al. 2017). It is sufficiently close to the RJ limit (ν^0) for our purposes. We apply a simple, first-order correction to obtain the relevant beam for the CMB blackbody spectrum using the band effective frequencies from Thornton et al. (2016).²¹ For this radiation with a band effective frequency ν_{CMB} , the

²¹ For the foreground modeling in Choi et al. (2020), these effective frequencies were re-computed with improved passband data and upgraded code (as described in Appendix C). Since the beam analysis for DR4 was done before this new work on the effective frequencies, the values from Thornton et al. (2016) were used. Considering the uncertainties on the passbands, the two sets of effective frequencies are consistent. In addition, given that the uncertainty on the beams is subdominant in the power spectrum analysis, whether one uses the effective frequencies from Thornton et al. (2016) or Choi et al. (2020) for the beam spectral correction does not have a significant effect on the results.

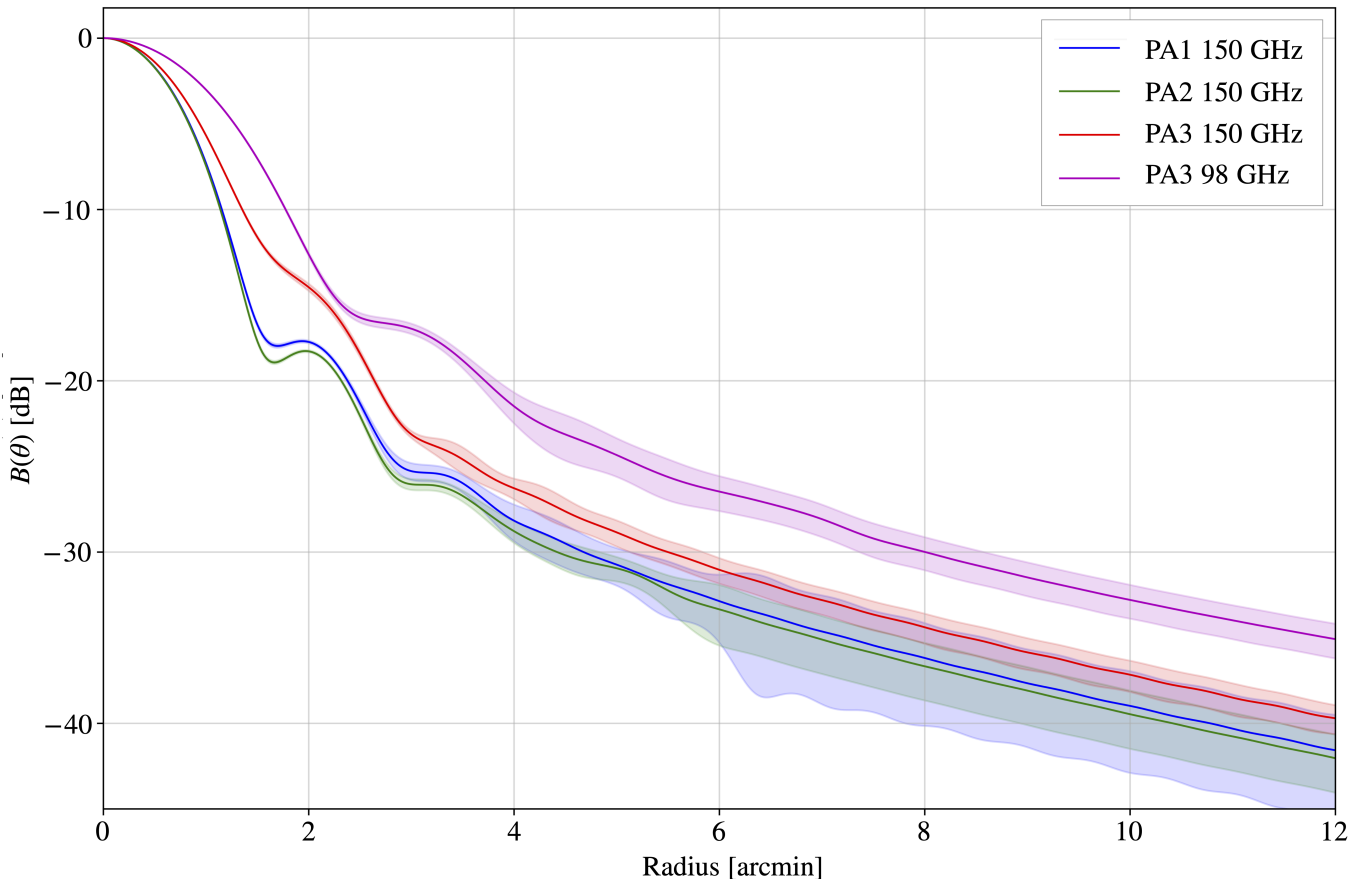


FIG. 8.— The estimated instantaneous beam profiles for S15, for the three arrays (PA1, PA2 and PA3), and for both 98 GHz and 150 GHz for PA3. The shaded bands indicate the 1σ uncertainty bounds. The uncertainties are strongly correlated between radial bins. The S13, S14 and S16 beam profiles are similar, and are included in the public data release. At 150 GHz, the beam for PA3 is quite different from those for PA1 and PA2. This is because PA3 was not focused as well, due to its position, further off-axis, compared to PA1 and PA2.

beam is taken to be $B'(\ell) = B(\ell\nu_{RJ}/\nu_{\text{CMB}})$, where ν_{RJ} is the effective frequency for radiation with a Rayleigh-Jeans spectrum. This is a $\lesssim 1.5\%$ effect for $\ell < 10,000$.

The resulting beam models are referred to as the “instantaneous” beams, and it is in fact these corrected beams that are shown in Figures 8 and 9.

4.5. Jitter Beams

In practice, the effective beam for a given sky region, season, detector array, and frequency is broader than the instantaneous beam. This is due to combining observations taken on multiple different nights throughout each season, so the resulting effective beam for each map is not as sharp as the beam inferred from planet observations which have been carefully recentered before co-adding. Broadening can be caused by variations in the pointing and global alignment, as well as possible small changes in the beam over the course of a season.

As described in Aiola et al. (2020), ACT’s blind pointing accuracy for DR4 is comparable to the average beam FWHM. If left uncorrected, this would significantly broaden the effective beams. Instead, as was done in Louis et al. (2017), we correct the pointing by comparing the observed positions of bright point sources to their known catalog positions. This is done for each 10-minute section of the time-ordered data from each detector array, as described in Aiola et al. (2020). Instead of performing

the fit in the time domain as in Louis et al. (2017), due to the larger data volume the fit is now performed in map-space. The resulting fit is obtained more quickly, but is slightly less accurate, leaving a larger residual variation in the beam due to pointing uncertainty.

This residual variation is captured by the effective beam, B_ℓ^{eff} , which is parametrized in terms of the instantaneous beam, B_ℓ , and a correction term that is Gaussian in ℓ , as

$$B_\ell^{\text{eff}} = B_\ell \times e^{-\ell(\ell+1)V/2}. \quad (22)$$

If we were to interpret the Gaussian correction term as arising purely from residual pointing errors, then V would be the residual pointing variance in square radians, which is why it is also referred to as “pointing jitter.” This variance, which in practice also includes the additional errors due to alignment and seasonal changes in the beam, is expected to be different for each sky region used in the power spectrum analysis. These different regions (Deep1, Deep5, Deep6, Deep56, Deep8, BOSS, and AdvACT) are shown in Figure 2 of Choi et al. (2020).

To estimate V , in each region we create a catalog of the brightest point sources that are found in the maps using a matched filter, that have a signal-to-noise of at least 10 in each map, and that are matched to catalogs of known sources. For each of these sources, we then obtain an estimate of the variance V and its associated

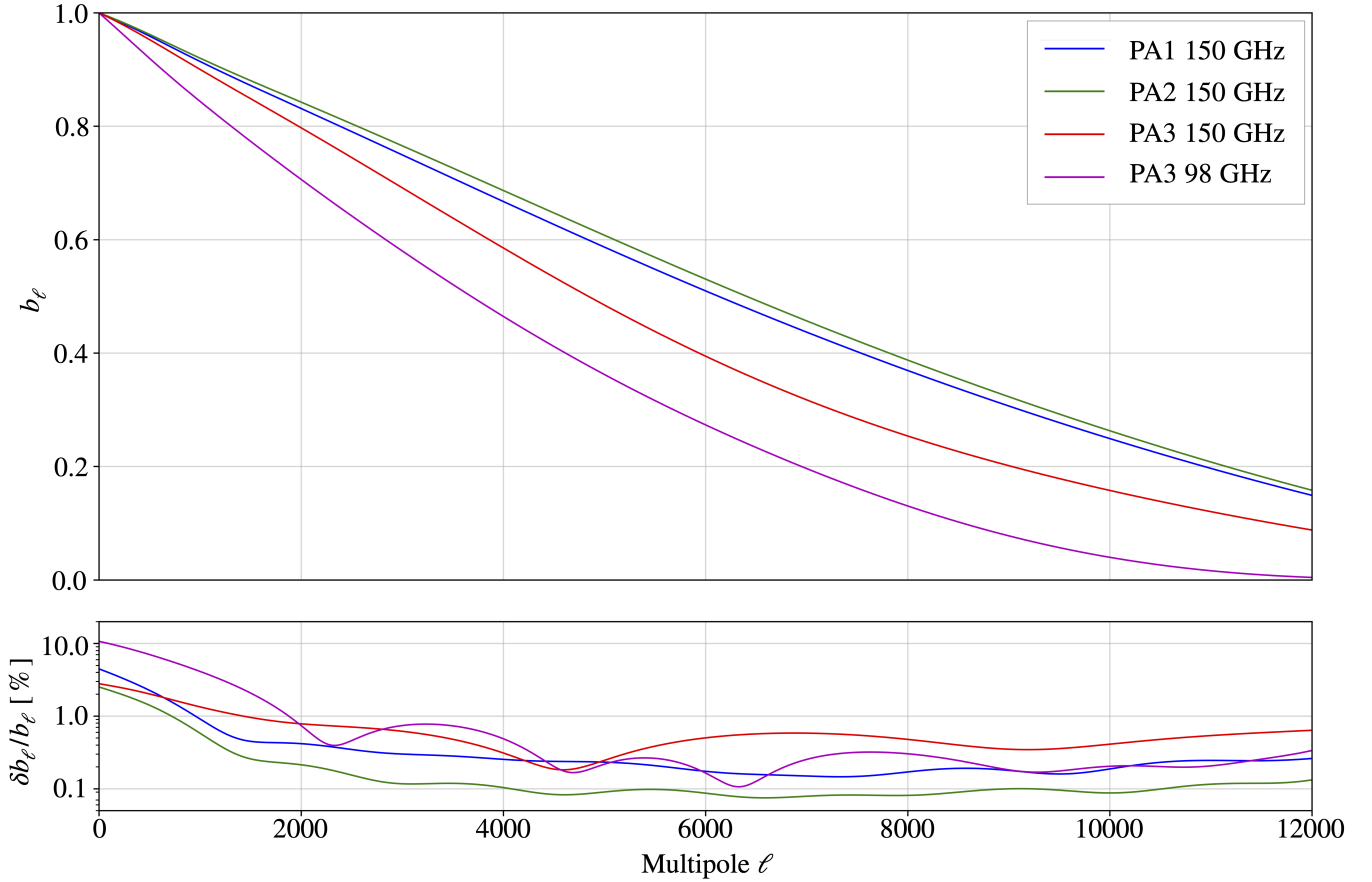


FIG. 9.— Instantaneous beam transforms and their uncertainties for the S15 data, for all the detector arrays. The uncertainties are strongly correlated between multipoles. For context when looking at this figure along with Figure 8, in the mapping from angle to multipole ($\ell \simeq \pi/\theta$), $1'$ corresponds to $\ell \simeq 10800$, $1.8'$ corresponds to $\ell \simeq 6000$, and $10.8'$ corresponds to $\ell \simeq 1000$.

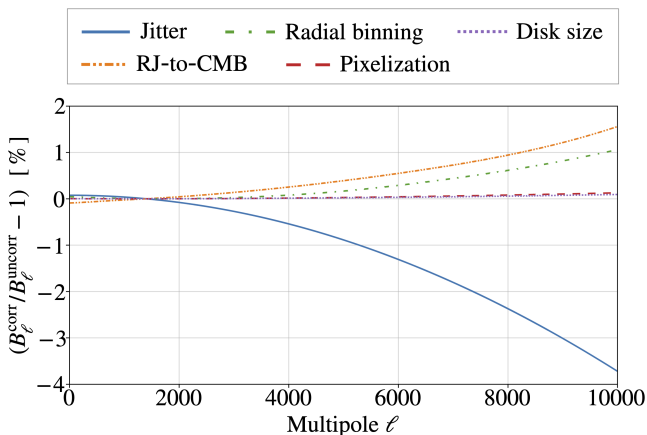


FIG. 10.— Change in the beam transform (in %) due to applying different corrections for S15 PA2 at 150 GHz. For this plot, the beam transforms have all been normalized at $\ell = 1400$. This corresponds roughly to our effective calibration scale, so any change in our beam at this multipole would be corrected by our subsequent calibration to *Planck*.

uncertainty for each season, array, and frequency. To do this, we crop a section ($10' \times 10'$) of each map around the source and remove large-scale variations by high-pass filtering the data (by multiplying the data by $1-G$, where G is a Gaussian filter with a FWHM of $8'$). We then

compute the effective beam using Equation 22, multiply this model by the local pixel-window of the cropped map, transform it into map space, high-pass filter it, and then compute the χ^2 using an estimate of the local white noise amplitude. An example of such an individual source fit is shown in Figure 11.

This procedure was validated by verifying that when simulated point sources convolved with Equation 22 (with a known value for V) are injected into real CMB data and run through the pipeline, the input value for V is recovered.

For convenience, we exclude sources identified as galaxies, pairs of galaxies, and planetary nebulae in the known source catalogs from this fitting procedure, as these types of sources sometimes appear to ACT to be extended. This leaves approximately 20 sources (for the smaller, deep regions) to 520 (for the largest region) that are included in the fits. They are primarily quasars and radio sources.

Once all N sources have been fit using this method, we use the resulting set of pointing jitter estimates \mathbf{d} and their uncertainties $\boldsymbol{\sigma}$ to obtain an estimate for the mean and intrinsic scatter, V and σ_V , of the effective pointing jitter in each map. The likelihood for this is written as

$$\mathcal{L}(\mathbf{d}, \boldsymbol{\sigma} | V, \sigma_V) = A(\boldsymbol{\sigma}, \sigma_V) e^{-\sum_{i=1}^N (d_i - V)^2 / (2(\sigma_i^2 + \sigma_V^2))}, \quad (23)$$

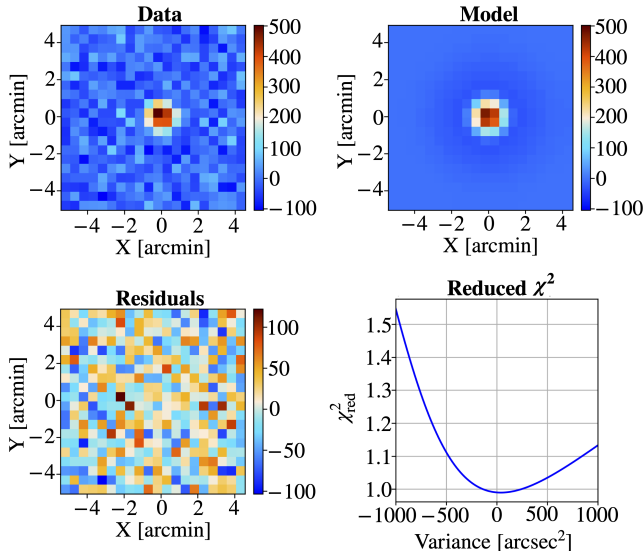


FIG. 11.— Example of the residual pointing variance, estimated by fitting an effective beam to a point source. The top panels compare the flux from the best-fitting model (right) to the data (left), with the residuals (bottom left), and the reduced χ^2 as a function of pointing variance (bottom right). In each case, the data is high-pass filtered, and the colorbars are in units of μK . The best-fit variance for this source is $41 \pm 93 \text{ arcsec}^2$.

where $A(\sigma, \sigma_V) = \Pi_i^N [2(\sigma_i^2 + \sigma_V^2)]^{-1/2}$ is a normalization factor. We estimate the posterior distribution for V and σ_V using MCMC with the Metropolis-Hastings algorithm (Hastings 1970), assuming a uniform prior on V . An example of these posterior distributions for one of the maps is shown in Figure 12.

This likelihood, which was not used for previous ACT analyses, better takes into account the intrinsic scatter in the residual pointing variance, resulting in an improved estimate.

After computing the pointing variance independently in each map, we use Equation 22 to obtain an effective beam, and associated covariance matrix, for each season, region, detector array, and frequency. An example of the effect of the pointing jitter correction on the beam transform is shown in Figure 10. The jitter values and uncertainties are given in Table 3 and the resulting solid angles and uncertainties for the effective beams are given in Table 4.

The jitter values for PA3 at 98 GHz are significantly greater than those at 150 GHz. This empirical finding serves as a useful reminder that this “jitter” encompasses not only pointing errors, but also any changes in the beam throughout each season. These changes may be greater at 98 GHz since that is where the beam is broader.

Some of the best-fit jitter values for the Deep8 region are negative, but considering their uncertainties, they are consistent with positive values. We allow for negative jitter values in the fits in order to account for all possible effects on the beams, not just those due to pointing. While we have not investigated whether these negative values are related to a data quality issue, it should be noted that the data from Deep8 were not used in the cosmological analysis due to the poor cross-linking in the region.

In Appendix C we provide the effective beam solid an-

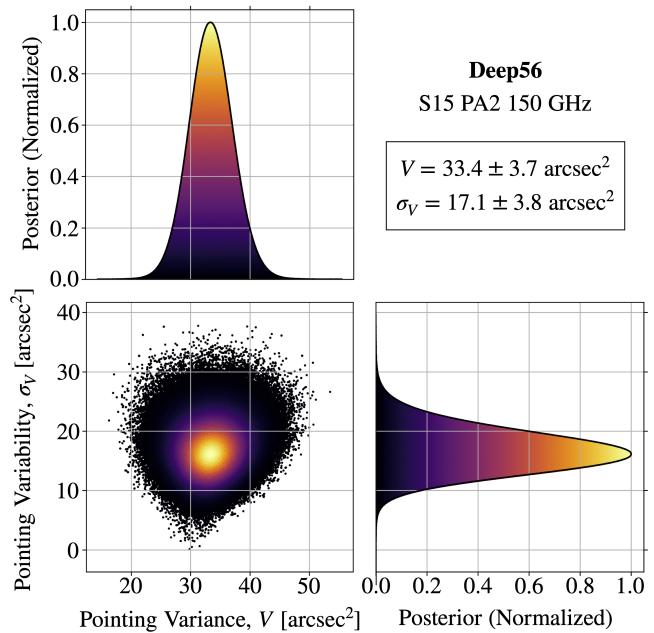


FIG. 12.— Distribution of the residual pointing variance, V , and pointing variability, σ_V , for the S15 PA2 beam at 150 GHz in the Deep56 region, resulting from fitting a pointing variance model to point sources. Notice here that the intrinsic scatter (measured with σ_V) is significantly larger than the measurement uncertainty on V .

gles for beams at different frequencies. We applied the first-order spectral correction from §4.4 with effective central frequencies corresponding to synchrotron emission, dust, and the thermal Sunyaev-Zel’dovich effect. These beam solid angles differ from the CMB estimates by 0.3 to 5.3%.

4.6. Beam Transform Covariance Matrix

There have been several modifications since Hasselfield et al. (2013) in how we estimate uncertainties in the beam model. Previously, the non-linear scaling parameter ℓ_{max} was varied in the fits, but the uncertainty in this parameter was not propagated to the covariance matrices. We now estimate the covariance of each beam transform, b_ℓ , by computing the Legendre transform of the sampled posterior distribution for the parameters describing the radial profile, as in Equation 21, and then applying the small-order corrections from §4.4. We then estimate the covariance matrix from this suite of ℓ -space beam transforms and the added uncertainty associated with the jitter correction. These matrices are large, since the beam transforms are computed at each integer ℓ from 0 to 30,000, so we do not store them in their entirety. Rather, we decompose each matrix into independent modes (via singular value decomposition) and store the largest 10 modes. We find that this is sufficient to capture the majority of the covariance (we do not discard any modes with a singular value larger than 10^{-3} of the maximum value).

We include additional modes to account for possible uncertainty due to variations in the surface rms ϵ (which enters into the calculation of the Ruze beam, as described in §4.2.5, and was fixed to our estimate of $20 \mu\text{m}$),²² as

²² We do not account for possible variations in the other mea-

TABLE 3
JITTER VALUES AND UNCERTAINTIES (ARCSEC²).

Array	Band	Season	Deep1	Deep5	Deep6	Deep56	Deep8	BOSS	AdvACT
PA1	150 GHz	S13	47.6 ± 18.5	26.6 ± 8.9	30.3 ± 10.2	-	-	-	-
		S14	-	-	-	12.1 ± 4.1	-	-	-
		S15	-	-	-	34.9 ± 4.4	-2.0 ± 15.4	30.8 ± 3.3	-
PA2	150 GHz	S14	-	-	-	22.5 ± 3.2	-	-	-
		S15	-	-	-	33.4 ± 3.7	4.2 ± 6.9	13.3 ± 3.3	-
		S16	-	-	-	-	-	-	35.7 ± 2.4
PA3	150 GHz	S15	-	-	-	6.6 ± 5.8	-21.3 ± 24.1	65.1 ± 6.8	-
		S16	-	-	-	-	-	-	12.2 ± 4.5
PA3	98 GHz	S15	-	-	-	190.9 ± 7.0	163.2 ± 18.5	246.0 ± 7.2	-
		S16	-	-	-	-	-	-	238.2 ± 5.8

TABLE 4
EFFECTIVE BEAM SOLID ANGLES AND UNCERTAINTIES (NSR).

Array	Band	Season	Deep1	Deep5	Deep6	Deep56	Deep8	BOSS	AdvACT
PA1	150 GHz	S13	206.1 ± 5.0	202.7 ± 4.1	203.3 ± 4.2	-	-	-	-
		S14	-	-	-	197.2 ± 3.4	-	-	-
		S15	-	-	-	199.0 ± 8.9	193.0 ± 9.0	198.3 ± 8.9	-
PA2	150 GHz	S14	-	-	-	183.9 ± 3.3	-	-	-
		S15	-	-	-	190.3 ± 4.8	185.7 ± 4.8	187.1 ± 4.7	-
		S16	-	-	-	-	-	-	188.2 ± 4.8
PA3	150 GHz	S15	-	-	-	267.1 ± 5.5	262.0 ± 6.9	278.0 ± 5.8	-
		S16	-	-	-	-	-	-	236.3 ± 8.5
PA3	98 GHz	S15	-	-	-	510.7 ± 22.0	506.0 ± 21.0	520.2 ± 22.4	-
		S16	-	-	-	-	-	-	490.3 ± 22.4

well as the range over which the model for each beam offset is fit (which initially was 3.5′–10.0′). For the surface rms, the values explored are 20 μm and 30 μm. For the region over which the offsets are fit, the three independent ranges explored are 3.5′–5.0′, 5.0′–7.0′, and 7.0′–10.0′. We store the top 3 modes associated with these model variations. The final uncertainty for each beam is thus composed of a total of 13 modes. These beam uncertainties are later added to the data covariance matrix as part of the power spectrum analysis pipeline.

An example of the effect of these additional modes on the beam transform uncertainties is shown in Figure 13. The significant increase in the uncertainty at low multipoles is mainly due to the inclusion of the different fit ranges for the offsets. Previously, as described in Hasselfield et al. (2013), the uncertainties were simply doubled from their formal values to account for potential systematic variations due to different fitting ranges. Even though this earlier method led to a smaller estimate of the beam uncertainties at low multipoles, this did not have a significant effect on our results for DR3, since the uncertainty on the beams is subdominant in the power spectrum analysis, and a low- ℓ cutoff of 500 (350) was applied to the TT (TE and EE) data.

Another difference compared to the DR3 analysis in Louis et al. (2017) is the treatment of calibration uncertainty. As described in Choi et al. (2020), for each season, region, array, and frequency the angular power spectra from ACT are calibrated to the *Planck* temper-

sured parameter for the Ruze beam, the correlation length c , since it is well constrained by our measurements.

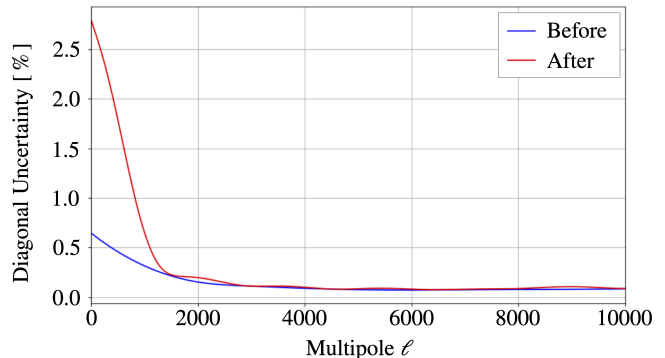


FIG. 13.— Uncertainties on the beam transform (in %) for S15 PA2 at 150 GHz, before (blue) and after (red) adding the additional modes to account for different surface rms values and different fit ranges for the beam offset levels. These uncertainties are included in Figure 9.

ature maps in the range $600 < \ell < 1800$. The Louis et al. (2017) analysis factored out the beam amplitude and uncertainty at an “effective” calibration scale (chosen in that case to be $\ell = 1400$).²³ However, we have changed this procedure for DR4 to reflect the fact that we are not calibrating the data at a single ℓ , but over a range of ℓ values. We now simply normalize the beam to unity at $\theta = 0$ and treat the calibration, and its uncertainty, separately in the ACT analysis. This is why the fractional beam uncertainties in Figure 9 are no longer

²³ This procedure is described by Equation 11 of Hasselfield et al. (2013).

smallest at $\ell = 1400$, differing from those reported for ACT DR3.

5. POLARIZATION

5.1. Main Beam

Measurements of Uranus in polarization at visible and near-infrared wavelengths have shown that its disk-integrated polarization is less than 0.05% (Schmid et al. 2006). While we are not aware of any similarly precise data at millimeter wavelengths, it is expected that the relevant scattering effects in the planetary atmosphere would be much weaker, resulting in net polarization levels considerably lower than the measurement cited above. Measurements by *Planck* at 100 and 143 GHz place 68% (95%) confidence upper limits on the polarization fraction of Uranus at 2.6% (3.6%) and 1.5% (2.0%), respectively (Planck Collaboration Int. LII et al. 2017).

Since we do not expect Uranus to be significantly polarized in the bands observed by ACT, we interpret any polarization response measured from Uranus as being due to temperature-to-polarization (T-to-P) leakage. Although this leakage is relatively small in magnitude, the ACT DR4 data are now sensitive enough that we must account for it in our analysis. To do this, we use observations of Uranus to build an ℓ -space T-to-P leakage function for each season, array, and frequency. This measurement and correction for the leakage in the main part of our beams is new to DR4 (see also Aiola et al. 2020; Choi et al. 2020).

We begin by making maps of the Q and U Stokes parameters for the same set of Uranus observations chosen in §4.1 to fit the main beam. We then convert each set of $\{Q, U\}$ polarization maps to the radial Stokes parameters $\{Q_r, U_r\}$ (following the third definition of cross polarization in Ludwig 1973) using the flat-sky approximation:

$$Q_r(\boldsymbol{\theta}) = Q(\boldsymbol{\theta}) \cos 2\phi_\theta + U(\boldsymbol{\theta}) \sin 2\phi_\theta \quad (24a)$$

$$U_r(\boldsymbol{\theta}) = U(\boldsymbol{\theta}) \cos 2\phi_\theta - Q(\boldsymbol{\theta}) \sin 2\phi_\theta, \quad (24b)$$

where $\boldsymbol{\theta} \equiv (\theta, \phi_\theta)$ are standard polar coordinates with the beam centroid as their origin and ϕ_θ increases clockwise from the positive y -axis (assuming that one uses the convention in which $+x$ points to the right and $+y$ points upward). Here Q and U follow the COSMO convention (Górski et al. 2005a), whereas for the ACT maps released as part of DR4, the polarization components are defined by the IAU convention (Hamaker & Bregman 1996). The polarization convention initially used for $\{Q, U\}$ does not matter once we have transformed to $\{Q_r, U_r\}$.

While it is not obvious how the polarized E or B beams ought to behave at larger radii, we do have some intuition about Q and U . For an unpolarized point source such as Uranus, any polarized signal will be the result of beam differences between the two axes of a polarimeter. Although the beams may be slightly different, due to e.g., differential ellipticity, we still expect each of them to decay radially as $1/\theta^3$ as they are part of the same diffraction-limited optical system. The Q and U maps are essentially just radially independent linear combinations of the various detector axes in an array, appropriately weighted by the map-maker, so the same asymptotic behavior should apply. This asymptotic behavior

would then hold true for the Q_r and U_r maps as well. This was confirmed with simulations where the beams for individual polarimeter axes were constructed using an “elliptified” version of the azimuthally averaged intensity beam for a given season and array. We thus use the same basis functions in the core and the α/θ^3 term in the wing to fit the beam in polarization as we did for the main beam in §4.2.6.

Since we are ultimately interested in how leakage manifests itself in the angular power spectra, we need to translate any polarized beam models of the azimuthally averaged radial profiles of Q_r and U_r , \tilde{Q}_r and \tilde{U}_r , to an ℓ -space representation of E and B . Conveniently, \tilde{Q}_r and \tilde{U}_r have a direct correspondence to the azimuthally averaged ℓ -space E and B transforms, $\tilde{E}(\ell)$ and $\tilde{B}(\ell)$. As shown in Appendix D, there exists a simple relation between these components, which in the flat-sky approximation takes the form of a second-order Hankel transform:

$$\{\tilde{E}(\ell), \tilde{B}(\ell)\} = -2\pi \int \{\tilde{Q}_r(\theta), \tilde{U}_r(\theta)\} J_2(\ell\theta) \theta d\theta. \quad (25)$$

From the set of Uranus maps in the $\{Q_r, U_r\}$ basis, examples of which are shown in Figure 14, we thus construct average radial profiles for each season, array, and frequency, and we fit them in a similar way as the main beam is fit in §4. While for the usual temperature beam fitting pipeline we fit an offset to each individual Uranus profile before taking an average, we do not do this in polarization since we expect the mapping transfer function in that case to be zero.²⁴ The only difference in the fitting procedure here is that we do not include a scattering term in the polarized beam model, since it is not expected to matter to first order, and it is unclear how it would vary for the different detector polarizations.²⁵ Examples of the radial profile model fits in polarization are shown in Figure 15.

As can be seen in Figures 14 and 15, while our model is a good fit to the Q_r and U_r radial profiles, there are significant, quadrupole-like residuals in the maps when our model is subtracted. It turns out that the leakage beams are far less azimuthally symmetric than the temperature beams (as can be seen by comparing Figures 7 and 14). The treatment of this leakage will be revisited for upcoming ACT beam analyses and may be improved upon, for example, by fitting a 2D model to the polarized beam profiles in order to properly capture the asymmetry.²⁶ In the meantime, the fitting done here is sufficient. The level of residuals seen in Figure 14 has an insignificant effect on the results of the power spectrum analysis for DR4 (see §6.2). The features visible in the residual maps can be expected to occur due to physical effects. For example, we simulated the T-to-P leakage for a point source due to both differential beam ellipticity between the two axes of a polarimeter and a polarization angle off-

²⁴ Even if the mapping transfer function were non-zero in polarization, it would be a sub-percent-level effect in the measurement of the leakage beam, which itself is percent-level in terms of our power spectrum analysis and results, so the effect would be insignificant.

²⁵ Again, this would be a sub-percent-level effect in the measurement of the percent-level leakage beams, so there would be no significant impact on our analysis and results.

²⁶ Work on fitting the beams in 2D is in progress.

set. The resulting maps of the simulated leakage beam in Q_r and U_r had strong quadrupole-like features. As shown in Figure 15, the azimuthal averages of the residuals in Q_r and U_r are close to zero, which is why the fits work well, despite the appearance of the residuals in the maps.

The Q_r and U_r radial profiles fits and their uncertainties are transformed to ℓ -space using a similar approach as for the temperature data, except using Equation 25 instead of the usual Legendre transform. The transforms for our example case are shown in Figure 16. These transforms do not include corrections for small systematic effects, as was done for the main beam fitting pipeline in §4.4. This is because we are ultimately interested in the ratios of transfer functions for the leakage beams, in which these ℓ -space corrections cancel out. To determine the polarization leakage beams, an example of which is shown in Figure 17, we divide the E and B transforms by their corresponding (uncorrected) T transform. For PA1 and PA2 the leakage values are within 1.5% (comparable to the leakage values for SPT-3G, Dutcher et al. 2021), whereas for PA3, the leakage increases around $\ell \simeq 4000$ –6000, and reaches over 6% (4%) at $\ell \sim 10,000$ at 150 GHz (98 GHz). We attribute the higher leakage for PA3 to an imperfectly optimized horn design in this first generation of multichroic polarimeters.

The top 10 modes from the leakage beam covariance matrices are stored. In addition, similar to the main beam analysis, two modes are added to account for variations in the model for the temperature beam.²⁷ The final leakage beam uncertainties are thus comprised of 12 modes.

The leakage beams were used in the DR4 power spectrum likelihood, as explained in §5.3.

5.2. Polarized Sidelobes

As described in Louis et al. (2017) and Aiola et al. (2020), we detect polarized sidelobes of the main ACT beams. Although weak in amplitude, these sidelobes cause noticeable T-to-P leakage. The sidelobes for PA1 and PA2 were shown in Louis et al. (2017). While PA3 was not used in that analysis, it was mentioned at the time that polarized sidelobes were not detected for PA3. However, using additional observations of Saturn to conduct a more thorough analysis, we have detected sidelobes in PA3 at 150 GHz, with an amplitude roughly 10% that of the sidelobes in PA1 and PA2. The general features of the sidelobes are common to all detector arrays; they consist of a group of compact lobes, each resembling a slightly elongated image of the main beam, with approximate four-fold symmetry, strongly polarized perpendicular to the radius from the beam center (which corresponds to $-Q_r$). This results in most of the leakage due to the sidelobes being from temperature into E -mode polarization rather than into B -mode polarization. These sidelobes are stable in time. We also observe that sidelobes from Saturn are only seen if Saturn lies in a focal plane’s field of view. As explained in Section 3.8 of Aiola et al. (2020), this is consistent with the sidelobes being due to an optical effect inside the receiver.

²⁷ We only consider variations in the range over which the beam offsets are fit (3.5’–5.0’, 5.0’–7.0’, 7.0’–10.0’). The Ruze beam is a smaller effect, so it is (safely) ignored here.

The sidelobes for PA1 and PA2 are at a distance of roughly 15’ from the beam centroid, with an additional set visible at 30’ for PA1. The sidelobes for PA3 are located approximately 30’ to 40’ from the beam centroid. This larger angular separation means the strongest T-to-P leakage occurs at $\ell \simeq 300$ for PA3 compared to $\ell \simeq 500$ for PA1 and PA2. Also, since the sidelobes only map to the sky when the main beam is also in the field of view, fewer detectors are affected by each sidelobe for PA3. We do not detect any sidelobes in PA3 at 98 GHz, which implies they must either be of significantly lower amplitude than those at 150 GHz, or in a different position. For PA1, PA2, and PA3, the amount of solid angle contained in these sidelobes is roughly 1.2%, 3.4%, and 0.15%, respectively, of that in the main beam.

Preliminary studies suggest that these sidelobes are due to diffraction caused by the arrays of metal elements that make up ACT’s optical filters. We note that similar filters (Ade et al. 2006) are also used for the South Pole Telescope (SPT), POLARBEAR, and BICEP/Keck (Padin et al. 2008; Arnold et al. 2010; Keating et al. 2003) at varying locations in the optical path.²⁸

If the sidelobes were indeed due to diffraction from the filters, we would expect them to occur at a larger radius for lower frequencies. Based on our calculations, at 98 GHz, the sidelobes would appear starting at a distance of approximately 47’ from the beam centroid. Since this is roughly the size of our field of view, the sidelobes would then map to the sky for only a small fraction of the detectors, the ones at the edges of the focal plane. This is consistent with the lack of detectable sidelobes at 98 GHz. We do not apply any sidelobe corrections to the PA3 data at 98 GHz.

As mentioned in Aiola et al. (2020), to study the sidelobes we use observations of Saturn. While Saturn’s brightness appears to induce a non-linear response near peak amplitude, it is useful for studying the relatively weak sidelobes. There are no issues due to non-linearity when observing the sidelobes. This is confirmed by the fact that the amplitude of the sidelobes seen by a detector is consistent, whether they are seen before or after the detector sees Saturn’s peak.

We apply the same treatment to the sidelobes for all detectors at 150 GHz. In short, we model the sidelobes as a sum of polarized, spatially shifted copies of the main beam and fit the amplitudes of the T , Q , and U components of each beam instance using maps of Saturn. We then use this model to deproject the sidelobes from the time-ordered data prior to map-making. The idea is to subtract the total flux in the sidelobes, even if their shapes are not exactly zeroed out in the maps. It can be shown that this removes the low- ℓ T-to-P leakage. Maps of the sidelobes are shown in Figure 18, along with the T-to-P leakage functions.

For PA1 and PA2, we re-use the sidelobe models from DR3, constructed using observations of Saturn from 2014. For PA3, which was not part of DR3, we use Saturn observations from 2015 to characterize the sidelobes

²⁸ This issue with ACT is related to the location of the filters in the optical path and the diffraction angle produced by the filters relative to the effective view passed by the system at this location. So other experiments may or may not see such an effect, depending on the details of their optical implementation.

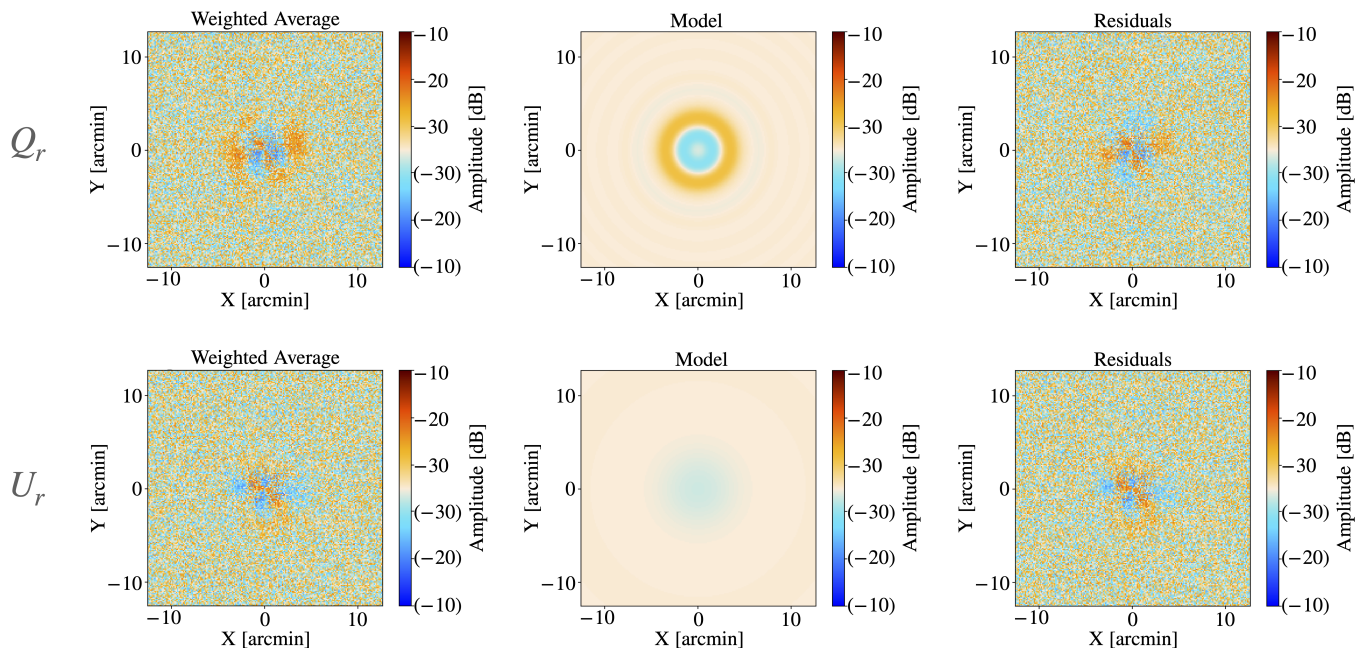


FIG. 14.— Maps of the leakage beam in Q_r (top) and U_r (bottom) for S15 PA2 at 150 GHz. In each case, the color scale is a symmetrical log scale in dB, with a linear threshold of -30 dB and negative values enclosed by parentheses. (Left) The weighted average of the Uranus observations used to characterize the beam. (Center) The model that is fit to the radial profile of the measured beam. (Right) The difference between the measured beam (left) and the beam model (center). Features in the residuals could arise due to physical effects, such as differential beam ellipticity between the two axes of a polarimeter or a polarization angle offset (Hu et al. 2003). The level of residuals seen in the last column has an insignificant effect on the results of the power spectrum analysis. As shown in Figure 15, the azimuthal averages of the residuals are close to zero, which is why the fits are successful.

for DR4 in the same way as was done for DR3.

Looking more closely at how these sidelobe models are constructed, we begin with a series of observations of Saturn to which we apply the same data selection criteria as we did to Uranus, as described in §2. We then map the chosen observations with the *moby2*²⁹ map-maker (the same map-maker that was used for the beam analysis in Louis et al. 2017) and coadd the maps together to produce one map of Saturn for each detector array (PA1, PA2, PA3). The maps are coadded with weights based on an estimate of their white noise level determined outside the planet region.

We then visually identify regions in the maps containing compact polarized sidelobes. As can be seen in Figure 18, the pattern of the sidelobes in each map has approximate four-fold symmetry, with four groups of sidelobes appearing at roughly equal distances from the beam centroid. For each sidelobe we choose how many copies of the main beam should be used to model it. Stronger, elongated sidelobes are modelled by two copies of the main beam, to capture the elongation; weaker sidelobes are modelled with a single copy of the beam. Then for each of the four groups of sidelobes we fit the position and amplitude of the copies of the main beam. In theory this fit could be done with any of the signals, but currently we perform this fit to a map of $P = \sqrt{Q^2 + U^2}$ because that is a bright, clear signal. With the model positions fixed, we then fit the amplitudes for each signal (T , Q_r , U_r) independently. The result is a base model for the sidelobes, but this model does not yet contain

per-detector detail.

We next account for the fact that the sidelobes do not always appear in all the detectors. This is relevant to the extent that the per-detector weights in the planet maps are different from the per-detector weights in the survey maps (which are used for the CMB analysis, for example). This could possibly be a significant (up to tens of percent) effect, since the *moby2* map-maker used to make the planet maps doesn’t weight the detectors by their noise, whereas the *enki* map-maker used for the survey maps does. The sidelobes occur primarily in the detectors at the periphery of the array, which is also where the noise tends to be higher, so this is a correlated effect.

To deproject the sidelobes from the per-detector time-ordered data, we need to estimate whether or not the sidelobes are seen by individual detectors, based on their position on the focal plane. The parameter we want to estimate is the radius of the “aperture”, or circle, around a detector such that if Saturn falls within the circle, the beam sidelobes are visible to that detector. To estimate this radius, we first divide the detectors into subsets. For PA3, the subsets are the three hexagonal wafers of detectors, described in Thornton et al. (2016). For each detector subset, we re-make maps of Saturn, and measure the total sidelobe flux in the resulting coadded map. We then compare these measurements to a model for the sidelobe flux as a function of radius, which scales with the fraction of detectors in each subset that would see each sidelobe.

As for the main beam analysis in §4.4, small corrections are made to the model to account for systematic effects.

The sidelobe removal described here is not perfect, so

²⁹ GitHub repository:
<https://github.com/ACTCollaboration/moby2>

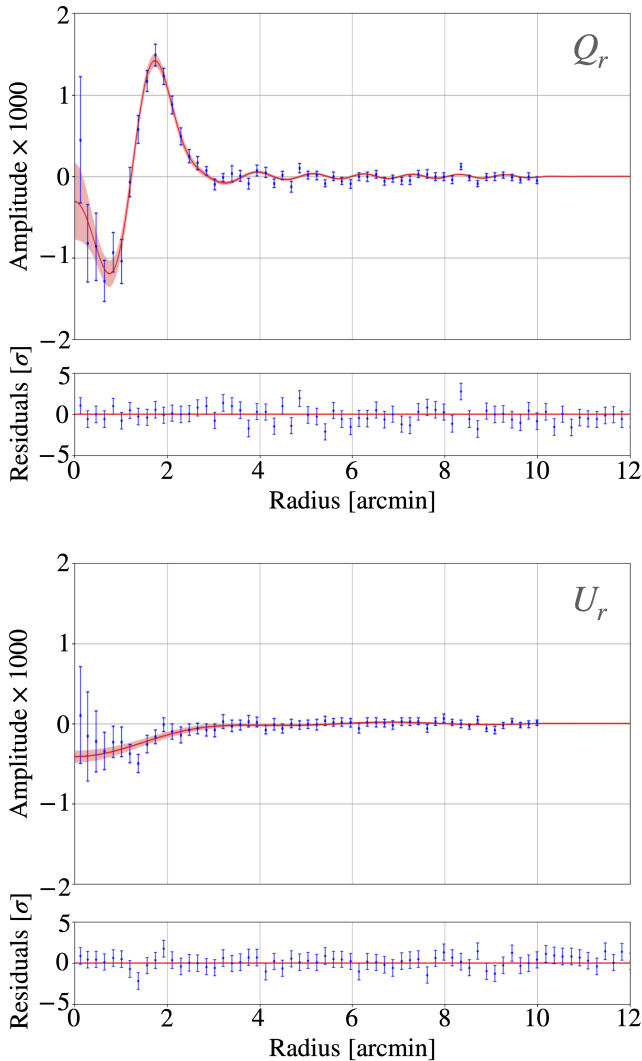


FIG. 15.— Measured radial polarization profiles (blue) in Q_r (top) and U_r (bottom) for S15 PA2 at 150 GHz and the model we fit to each profile (red), with the red shaded region indicating the 1σ model uncertainty bounds. The profiles have been normalized by the peak amplitude of the corresponding T beam (at the beam centroid, $\theta = 0$). Note that the bins are correlated.

there is still residual T-to-P leakage in the ACT data due to the sidelobes. The residual TE and TB leakages, shown in Figure 18, are estimated by making new maps of Saturn after our sidelobe removal. Since the sidelobes for PA3 are already significantly weaker than PA1 and PA2, we do not compute residuals for PA3. We add the residuals for PA1 and PA2 to the main beam leakage for use in the power spectrum analysis, as described in the next section.

The residual sidelobe leakage in Figure 18 can be compared to the leakage in the main beam shown in Figure 17. The effect of each of these components on the final spectra (if there is no leakage correction as in §5.3) is shown in Figure 19. At low ℓ , the residual sidelobe leakage dominates, whereas the main beam leakage grows larger by $\ell \sim 2000$, and dominates at high ℓ .

5.3. Leakage Correction

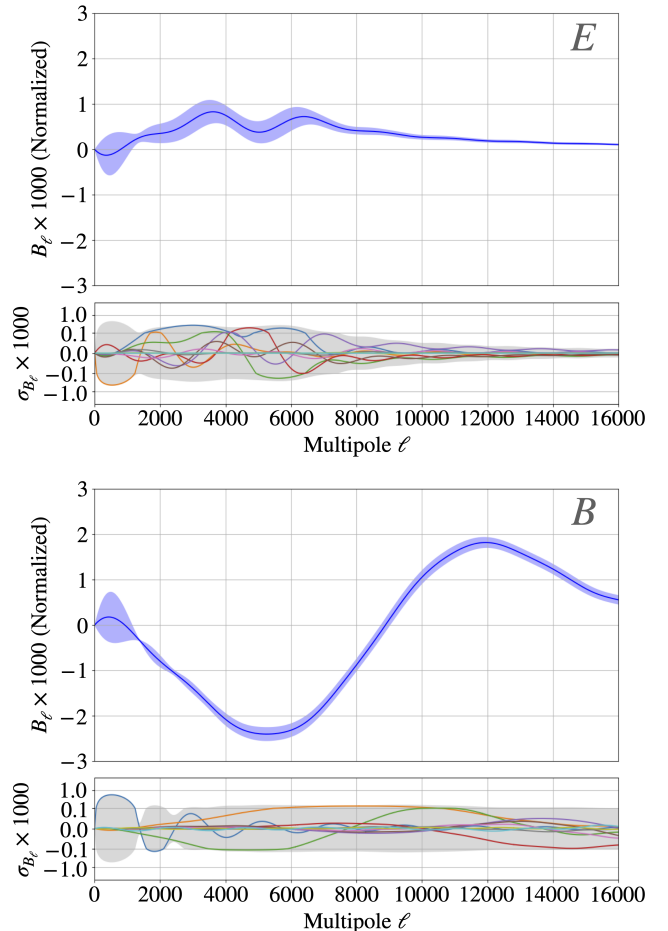


FIG. 16.— The transform of the radial beam profiles in E (top) and B (bottom) for S15 PA2 at 150 GHz. These transforms are normalized by the amplitude of the corresponding T beam transform at $\ell = 0$. The lower panels show the dominant independent modes of each transform’s covariance matrix (in color) and the magnitude of the diagonals (the gray shaded regions).

The main beam leakage and the residual leakage from the sidelobes are included in the power spectrum likelihood (see Section 12 of Choi et al. 2020) by making use of a leakage-corrected model for the TE and EE theory spectra, computed each time the likelihood is estimated. The corrected model spectra, $T_i E'_j$ and $E_i E'_j$, are related to the input theory spectra, $T_i T_j$, $T_i E_j$ and $E_i E_j$, via

$$T_i E'_j = T_i E_j + T_i T_j \gamma_j \quad (26a)$$

$$E_i E'_j = E_i E_j + T_i E_j \gamma_i + T_j E_i \gamma_j + T_i T_j \gamma_i \gamma_j. \quad (26b)$$

Here the i and j subscripts denote different spectra and the γ factors encode the ℓ -dependent leakage (the sum of the main beam leakage from §5.1 and the residual sidelobe leakage from §5.2) and are shown in Figure 20. At 150 (98) GHz, the amplitude of γ is never greater than 0.0035 (0.038). Around $\ell \sim 1000$ –3000, this leakage correction is roughly a few-percent (~ 1 –4%) effect for TE and a sub-percent (~ 0 –0.4%) effect for EE .

As explained in Choi et al. (2020), in the power spectrum analysis we first compute a power spectrum for each season, sky region, detector array, and frequency. We then coadd over seasons and arrays to obtain one spectrum per region and frequency. Finally, the regions are

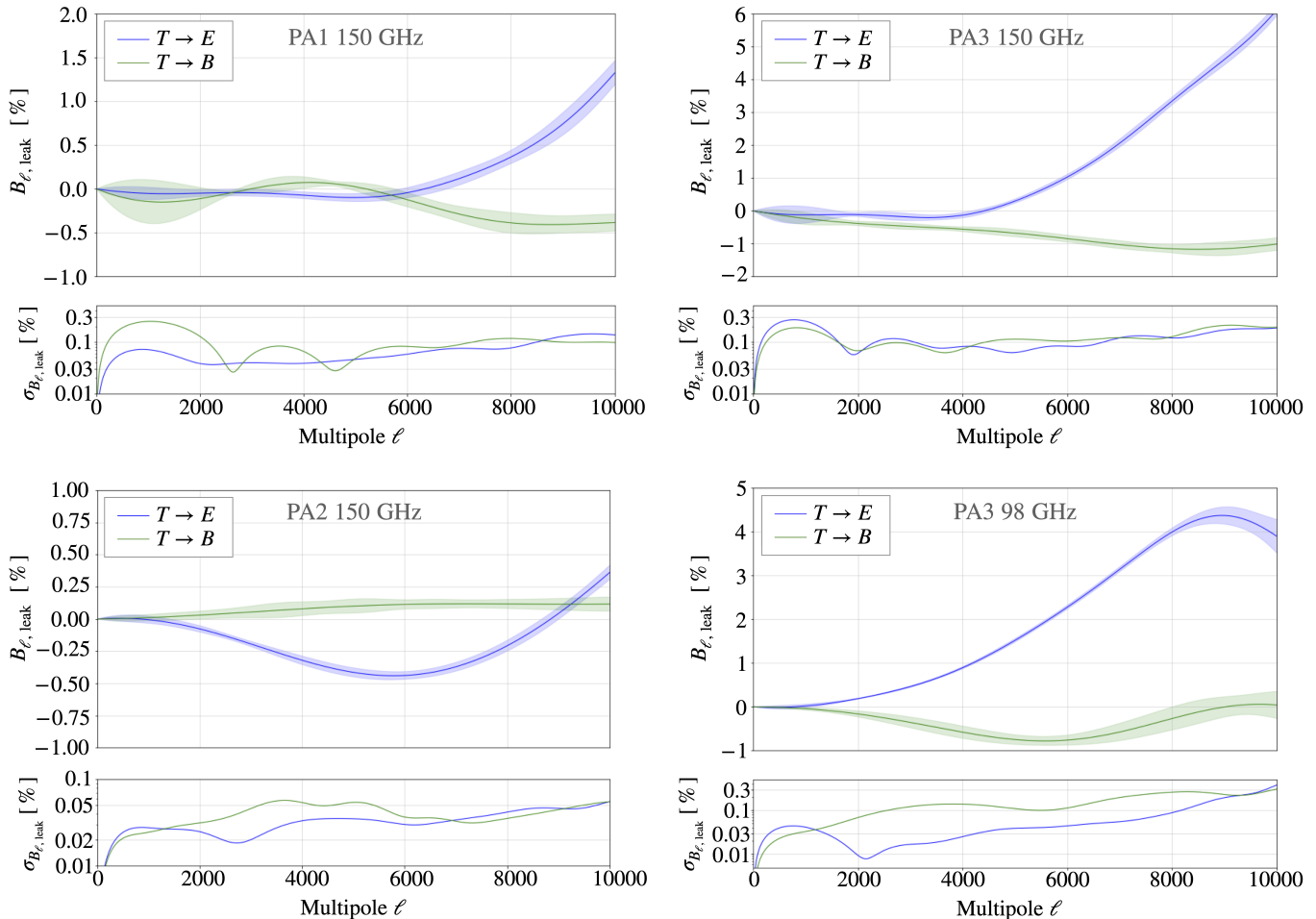


FIG. 17.— Measured temperature-to-polarization leakage in the main beams for S15 for all arrays, showing the temperature leaking into E -mode and B -mode polarization. Note the different y -axis ranges. For PA1 and PA2 the leakage values are within 1.5%, whereas for PA3, the leakage increases around $\ell \simeq 4000$ – 6000 , and reaches over 6% (4%) at $\ell \sim 10,000$ at 150 GHz (98 GHz). We attribute the higher leakage for PA3 to an imperfectly optimized horn design in this first generation of multichroic polarimeters. The uncertainties (lower panels) include the adjustments for model variability.

divided into two groups based on the detection thresholds for point sources: deep (Deep1, Deep5, Deep6, and Deep8) and wide (BOSS and AdvACT). The spectra for the regions in these two groups are coadded, resulting in a single deep and a single wide power spectrum at each frequency.

To obtain the γ factors, we coadd the leakage beams for individual seasons and detector arrays using the same weights used to coadd the spectra, giving an effective leakage beam for both the deep and wide coadded spectra at each frequency. The uncertainties in the γ factors are incorporated in the data covariance matrix by using the main leakage beam uncertainties and the sidelobe residuals to compute another covariance matrix (similar to the main temperature beam covariance matrix) which is then added (in quadrature) to the data covariance matrix.

As described in Aiola et al. (2020); Choi et al. (2020), including this leakage correction in the likelihood reduces the residuals compared to the best-fitting Λ CDM model but does not have a significant effect on the inferred cosmology. Choi et al. (2020) (in Section 12.3) also carried out a test by fitting for two scaling factors (one at

150 GHz and one at 98 GHz), that multiply the nominal values of the γ factors. The data support the baseline model, where the scaling factors are unity.

In the DR3 analysis in Louis et al. (2017) the residual leakage from the sidelobes was similarly treated as a systematic uncertainty in the cosmological power spectrum analysis, but the correction for the main beam polarized leakage is new to DR4.³⁰

6. DISCUSSION

6.1. Beam Products

As part of DR4, several ACT data products were made publicly available on the NASA Legacy Archive for Mi-

³⁰ As described in Choi et al. (2020), for DR4 the ACT team adopted a blinding strategy in an attempt to prevent confirmation bias on the cosmological parameters. The analysis pre-unblinding used maps that deprojected the polarized sidelobes, but did not include the additional polarized leakage correction from the main beam or the sidelobe residuals. The post-unblinding analysis revealed some features in the TE residuals to Λ CDM that led us to a more thorough search for possible sources of systematic uncertainty in the TE spectrum. As a result, it was at this point that we added the correction in the likelihood for the T-to-P leakage from both the polarized sidelobe residuals, and the main beam.

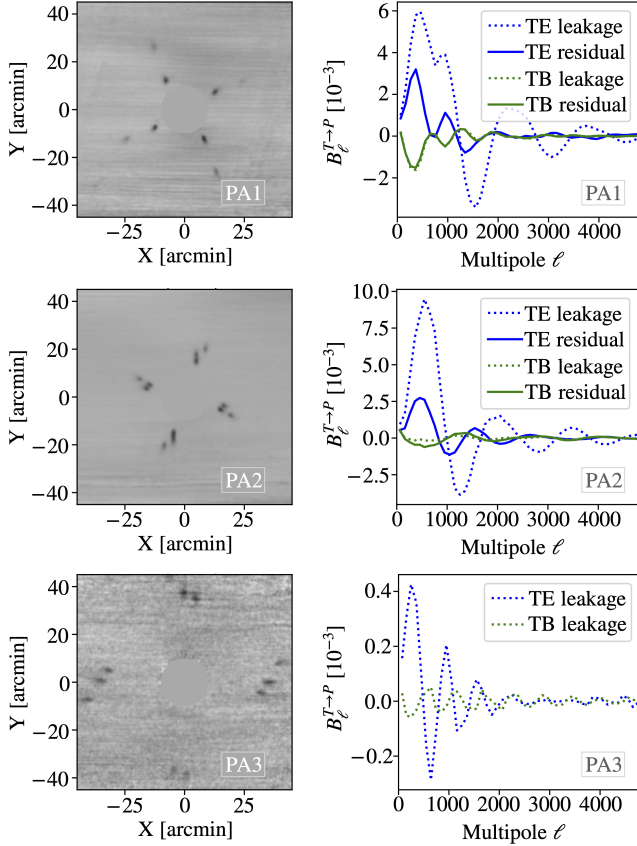


FIG. 18.— Sidelobes in the PA1, PA2, and PA3 detectors at 150 GHz. (Left) Maps of the polarized sidelobes, obtained by stacking observations of Saturn. The main part of the beam at the center of each image is masked. In each the grayscale is linear, indicating the sidelobe amplitude in the range from -0.002 (black) to $+0.001$ (white) relative to the main beam peak,^a with positive (negative) numbers corresponding to polarization parallel (perpendicular) to the radial direction. The complementary polarization leakage (corresponding to TB leakage) is smaller and not shown in the maps, but it is also estimated. (Right) The effect of the sidelobes on our measurements, expressed as beam transfer functions $B_{\ell}^{T \rightarrow E}$ and $B_{\ell}^{T \rightarrow B}$. Note that the scales on the right differ for all three arrays. The sidelobes are projected out of the time-ordered data prior to mapping, so the effective (residual) leakage is what affects our spectra. Since the sidelobes for PA3 are significantly weaker than PA1 and PA2, we do not compute residuals for PA3. No sidelobes are seen for PA3 at 98 GHz.

^aSaturn’s brightness appears to compress the detector gain by a few percent. The effect can be ignored here because it is a small correction on the sidelobe leakage, which itself is a small effect.

crowave Background Data Analysis³¹ (LAMBDA) and at the National Energy Research Scientific Computing Center (NERSC). Details of all these data products are given in [Mallaby-Kay et al. \(2021\)](#).

The beams described in this paper and used for the analysis in [Choi et al. \(2020\)](#) and [Aiola et al. \(2020\)](#) are included in the “ancillary products” section of the data release. For each season/region/array/frequency combination, there are multiple beam files. Both the real-space radial beam profiles and their harmonic-space transforms are provided, both for the instantaneous beams and for the beams with the jitter corrections included, and

³¹ lambda.gsfc.nasa.gov/product/act/actpol_prod_table.cfm

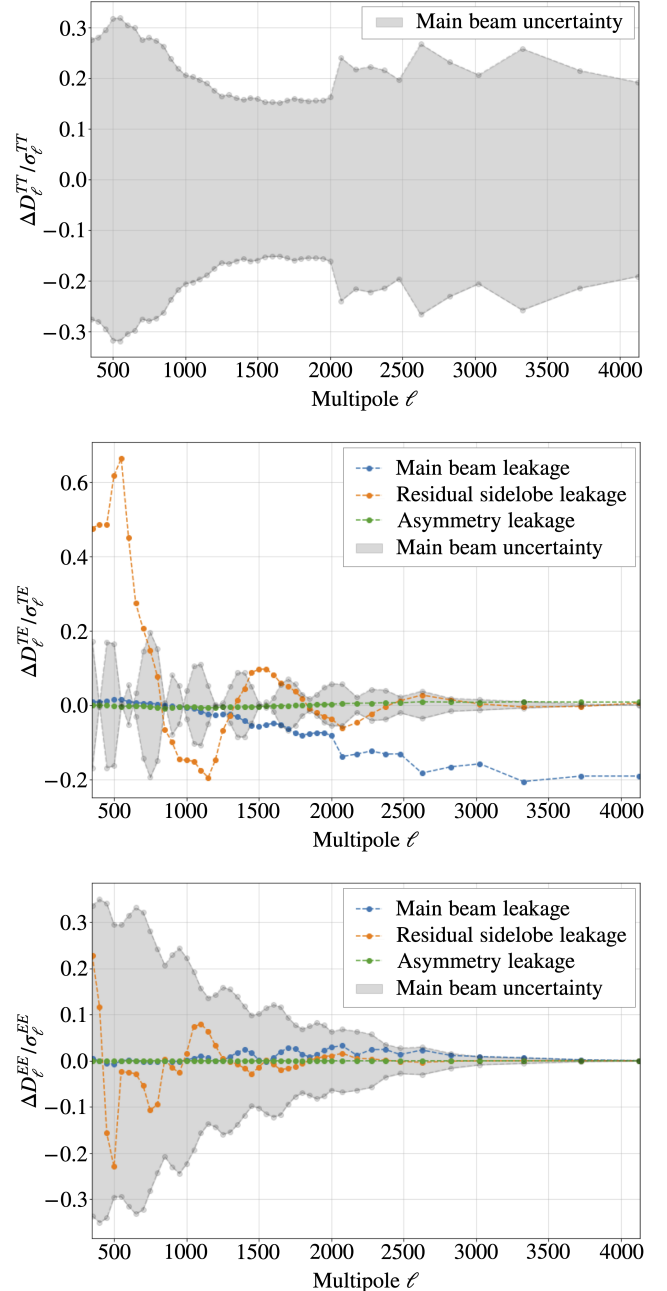


FIG. 19.— An example of the size of various effects on the inferred spectra, compared to our final uncertainties. Here we used the leakage in the main beam for S15 PA2 at 150 GHz, the residual sidelobe leakage for PA2 at 150 GHz, simulated leakage due to asymmetry for S15 PA2 at 150 GHz, and the uncertainty in the main beam for S15 PA2 deep56 at 150 GHz. The best fit Λ CDM plus foreground model for the deep 150×150 GHz spectra for ACT only from [Choi et al. \(2020\)](#) was used to compute the resulting change in D_{ℓ}^{TT} , D_{ℓ}^{TE} , and D_{ℓ}^{EE} due to these elements. These changes are then divided by the corresponding uncertainties σ_{ℓ}^{TT} , σ_{ℓ}^{TE} , and σ_{ℓ}^{EE} on the final binned deep spectra. At larger multipoles, eventually the uncertainty on the main beam would come to dominate. Note that we correct for the main beam leakage and the residual sidelobe leakage in the power spectrum likelihood, as explained in §5.3. The asymmetry leakage (simulated using an all-sky beam convolution code, as mentioned in §6.2) is negligible for all detector arrays. Sudden changes around $\ell = 2000$, as can be seen in the top panel, for example, are due to a change in the bin width at this scale.

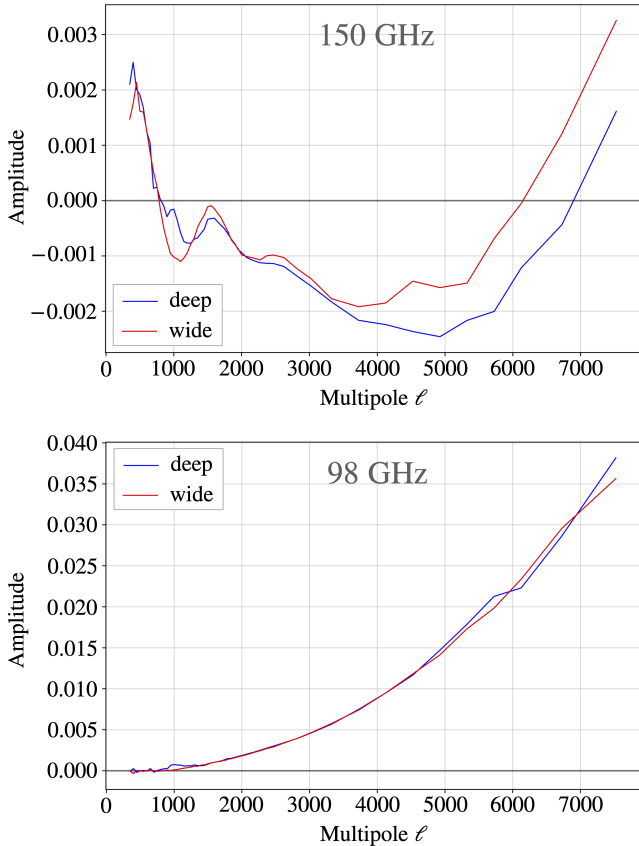


FIG. 20.— The γ factors encoding the ℓ -dependent T-to-P leakage, used in the power spectrum likelihood for both the deep and wide spectra at 150 GHz (top) and 98 GHz (bottom). The γ factors for 150 GHz include both the main beam leakage from §5.1 and the residual sidelobe leakage from §5.2. Since no sidelobes are seen at 98 GHz, those γ factors consist of only the main beam leakage.

in each case are available both with and without the Rayleigh-Jeans-to-CMB spectral correction.

Note that the instantaneous beams are not region-dependent, so only the jitter-corrected beams contain region flags in their filenames. The instantaneous beams are suitable for time-domain analysis, but they should not be used when analyzing maps. The maps released as part of DR4 have not been corrected for the instrument beam. When working with the maps, one should use the jitter-corrected beams where the pointing variance and its uncertainty have been accounted for. The beam harmonic-space transforms can be used to correct for the beam effects in harmonic space (Bond & Efstathiou 1987).

The TE and TB harmonic-space leakage beams for the main beam and the sidelobes and their residuals are also included in the ancillary products for DR4.

Since DR4 includes the data from DR3 as a subset, it is possible to compare some of the DR4 beams with the beams for the same season, sky region, detector array, and frequency released as part of DR3. As shown in Appendix E, despite the differences in the analyses, the beams from DR3 and DR4 are consistent.

6.2. Beam Asymmetry

The asymmetry of the ACT main beams is relatively well described as elliptical, with aspect ratios varying

from 1% to 20% across the different arrays, as shown in Table 1 of Choi et al. (2020). Here we give a brief summary of an investigation into this azimuthal asymmetry and the associated spurious signal.

Beam asymmetry is, to leading order, responsible for two effects. The asymmetric convolution creates anisotropy in the sky maps that distorts the shape of point sources and introduces statistical anisotropy in the inferred CMB. The magnitude of the effect is reduced by increased cross-linking: as the telescope observes a position on the sky using approximately orthogonal scan directions, the spurious signal from one scan roughly cancels with that of the other scan. The cross-linking in the ACT maps is sufficient to reduce the contamination to the power spectrum from this effect to an insignificant amount.

The second effect of beam asymmetry is to introduce T-to-P leakage. The beam asymmetry introduces a dependency in the time-ordered data on the position angle of the instrument: observations with different scan directions yield systematically different data. For asymmetric beams with a quadrupole shape, which is the dominant asymmetric azimuthal mode of our approximately elliptical beams, the dependence on the position angle is interpreted by the map-maker as a linearly polarized sky component (Hu et al. 2003). In contrast to the first effect, the T-to-P leakage is not averaged down by cross-linking. Instead, the leakage is reduced by the instrument’s orthogonally polarized co-pointing detectors: the leakage picked up by one detector approximately cancels with the leakage picked up by its partner.³² As mentioned in Choi et al. (2020) the residual leakage causes an additive bias to the TE and TB power spectra that is roughly constant with multipole with an amplitude that is less than 0.2σ away from zero. No attempt to remove this leakage has been made.

For the investigation described here, both effects were simulated in the time domain using an all-sky beam convolution code similar to Wandelt & Górski (2001); Reinecke et al. (2006); Prézeau & Reinecke (2010); Duivenvoorden et al. (2019).

6.3. Conclusion

In this paper, we have presented the analysis of the ACT beams for DR4, which includes data from 2013–16. Improvements to the beam pipeline include: better atmosphere subtraction for the Uranus maps, a new scattering term in the model that is fitted to the main beams, a better estimate of the uncertainty in these fits, and residual T-to-P leakage terms that are included in the ACT power spectrum likelihood. Considerable effort was spent developing a realistic model of the beams (including optical effects) and the mapping process, in order to study all elements of the analysis, including the propagation of systematic uncertainties.

The DR4 beams presented here were also used for the 2013–16 data that were part of the ACT DR5 maps (Naess et al. 2020). Finally, the pipeline presented here was used to obtain preliminary beams for the 2017–18

³² This cancellation is not perfect for detector pairs with incorrect relative gain or pointing or for pairs with slightly different beams, but these effects are small compared to the leakage from detectors without a partner, which see no reduction.

Advanced ACT data included in DR5. Looking forward, as we collect and analyze more ACT data and approach the cosmic variance limit, we will become increasingly sensitive to details of the instrument beams. While some details may change³³, we expect to adopt similar methods as described here for analysis of the post-2016 ACT data.

ACKNOWLEDGEMENTS

Support for ACT was through the U.S. National Science Foundation through awards AST-0408698, AST-0965625, and AST-1440226 for the ACT project, as well as awards PHY-0355328, PHY-0855887 and PHY-1214379. Funding was also provided by Princeton University, the University of Pennsylvania, and a Canada Foundation for Innovation (CFI) award to UBC. ACT operates in the Parque Astronómico Atacama in northern Chile under the auspices of the Agencia Nacional de Investigación y Desarrollo (ANID). The development of multichroic detectors and lenses was supported by NASA grants NNX13AE56G and NNX14AB58G. Detector research at NIST was supported by the NIST Innovations in Measurement Science program. Computations were performed on Tiger and Della as part of Princeton Research Computing resources at Princeton University, on Feynman at Princeton University, and on the Niagara supercomputer at the SciNet HPC Consortium. SciNet is funded by the CFI under the auspices of Compute Canada, the Government of Ontario, the Ontario Research Fund—Research Excellence, and the University of Toronto.

Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Industry Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

ML was supported by a Dicke Fellowship. ES and JD are supported through NSF grant AST-1814971 and AST-2108126. EC acknowledges support from the STFC Ernest Rutherford Fellowship ST/M004856/2, STFC Consolidated Grant ST/S00033X/ and from the Horizon 2020 ERC Starting Grant (Grant agreement No 849169). SKC acknowledges support from NSF award AST-2001866. JCH acknowledges support from NSF grant AST-2108536. ADH acknowledges support from the Sutton Family Chair in Science, Christianity and Cultures and from the Faculty of Arts and Science, University of Toronto. KM acknowledges support from the National Research Foundation of South Africa. LP gratefully acknowledges support from the Mishrahi and Wilkinson funds. ZX is supported by the Gordon and Betty Moore Foundation through grant GBMF5215 to the Massachusetts Institute of Technology.

We gratefully acknowledge the publicly available software packages that were used for parts of this analysis. They include `healpy` (Zonca et al. 2019), `HEALPix` (Górski et al. 2005b), `libsharp` (Reinecke & Seljebotn 2013), and `pixell`³⁴. This research made use of `Astropy`³⁵, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018). We also acknowledge use of the `matplotlib` (Hunter 2007) package and the Python Image Library for producing plots in this paper.

REFERENCES

- Abramowitz, M., & Stegun, I. A. 1972, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, tenth printing edn. (New York City: Dover), 364
- Ade, P. A. R., Pisano, G., Tucker, C., & Weaver, S. 2006, in *Millimeter and Submillimeter Detectors and Instrumentation for Astronomy III*, ed. J. Zmuidzinas, W. S. Holland, S. Withington, & W. D. Duncan, Vol. 6275, International Society for Optics and Photonics (SPIE), 248
- Aiola, S., Calabrese, E., Maurin, L., et al. 2020, *JCAP*, 2020, 047
- Akaike, H. 1973, Information theory and an extension of the maximum likelihood principle, in: B. N. Petrov and F. Csaki (eds), 2nd International Symposium on Information Theory (Budapest: Akadémiai Kiadó), 267-281
- Akaike, H. 1974, *IEEE Transactions on Automatic Control*, 19, 716
- Arnold, K., Ade, P. A. R., Anthony, A. E., et al. 2010, in *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy V*, ed. W. S. Holland & J. Zmuidzinas, Vol. 7741, International Society for Optics and Photonics (SPIE), 404
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Bond, J. R., & Efstathiou, G. 1987, *MNRAS*, 226, 655
- Born, M., & Wolf, E. 1980, *Principles of Optics*, 6th edn. (Oxford, England: Pergamon Press)
- Choi, S. K., Hasselfield, M., Ho, S.-P. P., et al. 2020, *JCAP*, 2020, 045
- Corkish, R. 1990, *IEEE Antennas and Propagation Magazine*, 32, 6
- ³³ Future changes likely will be informed in part by the work currently being done to model the beams in 2D.
- ³⁴ <https://github.com/simonsobs/pixell>
- ³⁵ <http://www.astropy.org>
- Darwish, O., Madhavacheril, M. S., Sherwin, B. D., et al. 2020, *MNRAS*, 500, 2250
- Duivendoorn, A. J., Gudmundsson, J. E., & Rahlin, A. r. S. 2019, *MNRAS*, 486, 5448
- Dutcher, D., Balkenhol, L., Ade, P., et al. 2021, *Phys. Rev. D*, 104
- Dünner, R., Hasselfield, M., Marriage, T. A., et al. 2013, *ApJ*, 762, 10
- Fowler, J. W., et al. 2007, *Appl. Opt.*, 46, 3444
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005a, *ApJ*, 622, 759
- . 2005b, *ApJ*, 622, 759
- Hamaker, J. P., & Bregman, J. D. 1996, *A&AS*, 117, 161
- Han, D., Sehgal, N., MacInnis, A., et al. 2021, *JCAP*, 2021, 031
- Hasselfield, M., Moodley, K., Bond, J. R., et al. 2013, *ApJS*, 209, 17
- Hastings, W. K. 1970, *Biometrika*, 57, 97
- Hincks, A. D., Acquaviva, V., Ade, P. A. R., et al. 2010, *ApJS*, 191, 423
- Hu, W., Hedman, M. M., & Zaldarriaga, M. 2003, *Phys. Rev. D*, 67
- Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90
- Hurvich, C., & Tsai, C. 1989, *Biometrika*, 76, 297
- Keating, B. G., Ade, P. A. R., Bock, J. J., et al. 2003, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4843, *Polarimetry in Astronomy*, ed. S. Fineschi, 284
- Keller, J. 1956, *IEEE Transactions on Antennas and Propagation*, 4, 312
- . 1959, *Journal of Applied Physics*, 30, 1452
- . 1962, *Journals of the Optical Society of America*, 52, 116
- Ledoit, O., & Wolf, M. 2003, *Journal of Empirical Finance*, 10, 603
- Louis, T., Grace, E., Hasselfield, M., et al. 2017, *JCAP*, 2017, 031
- Ludwig, A. 1973, *IEEE Transactions on Antennas and Propagation*, 21, 116

- Madhavacheril, M. S., Hill, J. C., Naess, S., et al. 2020, *Phys. Rev. D*, **102**
- Mallaby-Kay, M., Atkins, Z., Aiola, S., et al. 2021, *ApJS*, **255**, 11
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *The Journal of Chemical Physics*, **21**, 1087
- Morris, T. W., Bustos, R., Calabrese, E., et al. 2021, The Atacama Cosmology Telescope: Modeling Bulk Atmospheric Motion, [arXiv:2111.01319 \[astro-ph.IM\]](https://arxiv.org/abs/2111.01319)
- Naess, S., Hasselfield, M., McMahon, J., et al. 2014, *JCAP*, **2014**, 007
- Naess, S., Aiola, S., Austermann, J. E., et al. 2020, *JCAP*, **2020**, 046–046
- Naess, S. K. 2019, *JCAP*, **2019**, 060–060
- Namikawa, T., Guan, Y., Darwish, O., et al. 2020, *Phys. Rev. D*, **101**
- Padin, S., Staniszewski, Z., Keisler, R., et al. 2008, *Appl. Opt.*, **47**, 4418
- Page, L., Jackson, C., Barnes, C., et al. 2003, *ApJ*, **585**, 566–586
- Piazzo, L. 2017, *IEEE Transactions on Image Processing*, **26**, 5232
- Planck Collaboration Int. LII, Akrami, Y., Ashdown, M., Aumont, J., et al. 2017, *A&A*, **607**, A122
- Pope, A. C., & Szapudi, I. 2008, *MNRAS*, **389**, 766–774
- Prézeau, G., & Reinecke, M. 2010, *ApJS*, **190**, 267
- Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, *AJ*, **156**, 123
- Reinecke, M., Dolag, K., Hell, R., Bartelmann, M., & Enßlin, T. A. 2006, *A&A*, **445**, 373
- Reinecke, M., & Seljebotn, D. S. 2013, *A&A*, **554**, A112
- Ruze, J. 1966, *Proceedings of the IEEE*, **54**, 633
- Schmid, H. M., Joos, F., & Tschan, D. 2006, *A&A*, **452**, 657
- Schäfer, J., & Strimmer, K. 2005, *Statistical Applications in Genetics and Molecular Biology*, **4**
- Thornton, R. J., Ade, P. A. R., Aiola, S., et al. 2016, *ApJS*, **227**, 21
- Wandelt, B. D., & Górski, K. M. 2001, *Phys. Rev. D*, **63**, 123002
- Wright, E. L. 1976, *ApJ*, **210**, 250
- Zonca, A., Singer, L., Lenz, D., et al. 2019, *Journal of Open Source Software*, **4**, 1298

APPENDIX

A. SHRINKING ALGORITHM FOR RADIAL PROFILE COVARIANCE MATRIX

In general, shrinkage is useful when estimating a covariance matrix (or any matrix, really) from a limited number of data points (Schäfer & Strimmer 2005). This technique works by combining an empirical estimate of the covariance matrix (a high-dimensional estimate of the underlying covariance with little or no bias) with a model (a low-dimensional estimate which may be biased but has much smaller variance) to minimize the total mean squared error (sum of bias squared and variance) with respect to the true underlying covariance. This is useful when the off-diagonal elements of the covariance matrix are excessively noisy. One can analytically calculate the optimal combination of the low and high dimensional estimates, parametrized by the shrinkage intensity. A review of covariance matrix shrinkage with an example of application to cosmological analysis is given in Pope & Szapudi (2008). Shrinkage can result in a much better estimation of the covariance matrix when few measurements are available, and it does not adversely affect the covariance in the case of a large number of measurements. Since the method is computationally very simple, Pope & Szapudi (2008) suggest it should always be employed.

The method works as follows. Suppose we have n sets of observations and each observation yields a data vector \mathbf{x} of length p . In the case of the beams analysis for DR4, for a given season, detector array, and frequency, the number of observations n is between 6 and 51 (as listed in Table 1) and in each case the resulting data vector \mathbf{x} (radial profile) is of length $p = 55$. Let $x_i^{(k)}$ then represent the i^{th} element of the vector for the k^{th} observation. The estimated empirical mean of the i^{th} element across all observations is then $\bar{x}_i = (1/n) \sum_{k=1}^n x_i^{(k)}$.

If we define

$$W_{ij}^{(k)} = (x_i^{(k)} - \bar{x}_i)(x_j^{(k)} - \bar{x}_j), \quad (\text{A1})$$

and

$$\bar{W}_{ij} = \frac{1}{n} \sum_{k=1}^n W_{ij}^{(k)}, \quad (\text{A2})$$

then an unbiased, empirical estimate of the covariance, \mathbf{S} , of the data, is

$$S_{ij} = \widehat{\text{Cov}}(x_i, x_j) \quad (\text{A3a})$$

$$= \frac{n}{n-1} \bar{W}_{ij} \quad (\text{A3b})$$

$$= \frac{1}{n-1} \sum_{k=1}^n (x_i^{(k)} - \bar{x}_i)(x_j^{(k)} - \bar{x}_j). \quad (\text{A3c})$$

We can also compute the covariance of the elements of this covariance matrix,

$$\widehat{\text{Cov}}(S_{ij}, S_{lm}) = \frac{n}{(n-1)^3} \sum_{k=1}^n (W_{ij}^{(k)} - \bar{W}_{ij})(W_{lm}^{(k)} - \bar{W}_{lm}), \quad (\text{A4})$$

writing the variance of an individual covariance matrix entry as $\widehat{\text{Var}}(S_{ij}) = \widehat{\text{Cov}}(S_{ij}, S_{ij})$.

Let \mathbf{T} be the target matrix, our model with fewer (or no) free parameters. An example of such a matrix could be the identity matrix times a constant, or the target we use, which is described later, in Equation A7. Then we can similarly write an equation for $\widehat{\text{Cov}}(T_{ij}, S_{ij})$. The optimal shrinkage intensity, λ^* , and resulting final estimate of the covariance matrix \mathbf{C} are then given by

$$\hat{\lambda}^* = \frac{\sum_{i,j} \widehat{\text{Var}}(S_{ij}) - \widehat{\text{Cov}}(T_{ij}, S_{ij})}{\sum_{i,j} (T_{ij} - S_{ij})^2}, \quad (\text{A5})$$

$$\mathbf{C} = \hat{\lambda}^* \mathbf{T} + (1 - \hat{\lambda}^*) \mathbf{S}. \quad (\text{A6})$$

The expression for $\hat{\lambda}^*$ above is the practical estimator suggested by Schäfer & Strimmer (2005) in order to estimate the optimal shrinkage intensity. This is based on the analytic solution for the optimal shrinkage intensity, λ^* , introduced by Ledoit & Wolf (2003), which is identical to the expression above except the unbiased sample estimates $\widehat{\text{Var}}$ and $\widehat{\text{Cov}}$ are replaced by the true underlying Var and Cov . Before this analytic solution, matrix shrinkage was much less practical, since complex and computationally intensive methods were needed to find the optimal shrinkage intensity.

Looking more closely at Equation A5, the term $\widehat{\text{Var}}(S_{ij})$ in the numerator means that as the variances of the elements of the empirical covariance matrix decrease (so, as n becomes much larger than p) the shrinkage intensity decreases and our final estimate \mathbf{C} of the covariance matrix approaches the empirical estimate \mathbf{S} . In the case of our radial beam

profile, $n \ll p$, which is precisely when shrinkage is the most important. The term $\widehat{\text{Cov}}(T_{ij}, S_{ij})$ in the numerator accounts for the fact that \mathbf{S} and \mathbf{T} are estimated from the same data, so if some elements of \mathbf{T} are equal to elements of \mathbf{S} then the two terms in the numerator cancel for those elements, and they do not affect the estimate of $\hat{\lambda}^*$. The denominator in Equation A5 ensures that if our choice of \mathbf{T} is very different from \mathbf{S} , then $\hat{\lambda}^*$ will be small and \mathbf{C} will be close to \mathbf{S} . A poor choice of target matrix should not therefore negatively affect \mathbf{C} .

If Equation A5 leads to a value of $\hat{\lambda}^*$ that is greater than one, then we set $\hat{\lambda}^* = 1$, which means our final estimate of the covariance matrix is composed of only the target matrix. On the other hand, if the equation leads to a value of $\hat{\lambda}^*$ that is less than zero, we set $\hat{\lambda}^* = 0$, and our final estimate of the covariance matrix is equal to the empirical covariance matrix.

Several common choices for the target matrix are listed in Table 2 of Schäfer & Strimmer (2005), along with simplified expressions for the associated shrinkage intensities $\hat{\lambda}^*$. In this analysis we use a “diagonal, unequal variance” target matrix, that is, a matrix whose diagonal elements are equal to the diagonal elements of our empirical estimate \mathbf{S} of the covariance matrix and whose off-diagonal elements are zero:

$$T_{ij} = \begin{cases} S_{ii} & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \quad (\text{A7})$$

In this case the expression for the shrinkage intensity simplifies to

$$\hat{\lambda}^* = \frac{\sum_{i \neq j} \widehat{\text{Var}}(S_{ij})}{\sum_{i \neq j} S_{ij}^2}. \quad (\text{A8})$$

B. COMPUTING THE HARMONIC TRANSFORM OF THE BEAM

The spherical harmonics are functions defined on the sphere as

$$Y_l^m(\theta, \phi) = P_l^m(\cos \theta) \exp\{im\phi\}, \quad (\text{B1})$$

where P_l^m are the Legendre polynomials normalized for spherical harmonics. These are defined as

$$P_l^m(x) = (-1)^m \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l-|m|)!}{(l+|m|)!}} (1-x^2)^{|m|/2} \frac{d^{|m|}}{dx^{|m|}} P_l(x), \quad (\text{B2})$$

where the Legendre polynomials $P_l(x)$ are defined via the recurrence relation

$$P_0(x) = 1 \quad (\text{B3a})$$

$$P_1(x) = x \quad (\text{B3b})$$

$$nP_n(x) = (2n-1)xP_{n-1}(x) - (n-1)P_{n-2}(x). \quad (\text{B3c})$$

The spherical harmonics are orthonormal, so they obey the relation

$$\int_0^{2\pi} \int_0^\pi Y_l^m(\theta, \phi) \overline{Y_{l'}^{m'}}(\theta, \phi) \sin \theta \, d\theta \, d\phi = \delta_{ll'} \delta_{mm'}, \quad (\text{B4})$$

where \bar{z} is the complex conjugate of z and δ_{ij} is the Kronecker symbol.

In a spherical harmonic transform, we compute the coefficients f_l^m used to express a function $f(\theta, \phi)$ as

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_l^m Y_l^m(\theta, \phi). \quad (\text{B5})$$

The coefficients can be computed using the equation

$$f_l^m = \int_0^{2\pi} \int_0^\pi f(\theta, \phi) \overline{Y_l^m}(\theta, \phi) \sin \theta \, d\theta \, d\phi \quad (\text{B6a})$$

$$= \int_0^{2\pi} \int_0^\pi f(\theta, \phi) P_l^m(\cos \theta) \exp\{-im\phi\} \sin \theta \, d\theta \, d\phi. \quad (\text{B6b})$$

If $f(\theta, \phi)$ is independent of ϕ (as is the case for our beam), then we can write $f(\theta, \phi) = f(\theta)$ and the equation above becomes

$$f_l^m = \int_0^{2\pi} \exp\{-im\phi\} d\phi \int_0^\pi f(\theta) P_l^m(\cos \theta) \sin \theta \, d\theta. \quad (\text{B7})$$

The integral over ϕ then simplifies to

$$\int_0^{2\pi} e^{-im\phi} d\phi = 2\pi \delta_{m0}. \quad (\text{B8})$$

So f_l^m is only non-zero for $m = 0$, in which case we have

$$f_l^0 = 2\pi \int_0^\pi f(\theta) P_l(\cos \theta) \sin \theta \, d\theta \quad (\text{B9a})$$

$$= 2\pi \int_{-1}^1 f(\theta) P_l(\cos \theta) d \cos \theta. \quad (\text{B9b})$$

This is the equation for the Legendre polynomial transform, presented as a means of converting the radial beam profile $B(\theta)$ to the harmonic transform B_ℓ . However, this can be time-consuming to compute. For small beams such as ours, it is not necessary to work in the curved sky regime. We instead perform a 2D Fourier transform, which effectively becomes a Hankel transform, as shown below. The difference between the Hankel and Legendre polynomial transforms is less than a factor of 4×10^{-5} between $\ell = 0$ and $\ell = 10,000$ and the Hankel transform is much faster to compute.

Now let's consider the 2D Fourier transform of a function $f(x, y)$,

$$F(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp\{-i(xk_x + yk_y)\} dx dy. \quad (\text{B10})$$

Introducing the polar coordinates

$$\begin{aligned} x &= \theta \cos \phi & y &= \theta \sin \phi \\ k_x &= k \cos \psi & k_y &= k \sin \psi \end{aligned}$$

where θ and ϕ here correspond to the θ and ϕ in spherical coordinates used throughout the paper, we then have, in the flat sky approximation,

$$F(k \cos \psi, k \sin \psi) \equiv \mathcal{F}(k, \psi) = \int_0^\infty \int_0^{2\pi} f(\theta, \phi) \exp\{-i\theta k(\cos \phi \cos \psi + \sin \phi \sin \psi)\} \theta d\theta d\phi. \quad (\text{B11})$$

If our function is circularly symmetric, so independent of ϕ (as is the case for our beam model), we have $f(x, y) = f(\theta, \phi) = f(\theta)$ and the equation above becomes

$$\mathcal{F}(k, \psi) = \int_0^\infty \theta f(\theta) \int_0^{2\pi} \exp\{-i\theta k(\cos \phi \cos \psi + \sin \phi \sin \psi)\} d\theta d\phi \quad (\text{B12a})$$

$$= \int_0^\infty \theta f(\theta) \int_0^{2\pi} \exp\{-i\theta k \cos(\phi - \psi)\} d\theta d\phi \quad (\text{B12b})$$

$$= \int_0^\infty \theta f(\theta) \int_0^{2\pi} \exp\{-i\theta k \cos \alpha\} d\theta d\alpha \quad (\text{B12c})$$

$$= \int_0^\infty \theta f(\theta) 2 \int_0^\pi \exp\{-i\theta k \cos \alpha\} d\theta d\alpha. \quad (\text{B12d})$$

Using the integral representation

$$J_n(z) = \frac{(-i)^n}{\pi} \int_0^\pi \exp\{iz \cos \varphi\} \cos(n\varphi) d\varphi \quad (\text{B13})$$

for the Bessel functions J_n of the first kind, we have

$$J_0(z) = \frac{1}{\pi} \int_0^\pi \exp\{iz \cos \varphi\} d\varphi, \quad (\text{B14})$$

and so the final expression for the 2D Fourier transform of a circularly symmetric function $f(\theta)$ may be written as

$$\mathcal{F}(k) = 2\pi \int_0^\infty \theta f(\theta) J_0(-\theta k) d\theta \quad (\text{B15a})$$

$$= 2\pi \int_0^\infty \theta f(\theta) J_0(\theta k) d\theta, \quad (\text{B15b})$$

which is a Hankel transform of order zero, and where the last line follows from the identity $J_n(-z) = J_n(z)$ for integer n .

In order to compute the harmonic transform of our beam profile, we evaluate the expression above separately for the three main terms in our beam profile fit: the core term (composed of the sum of basis functions), the scattering term, and the $1/\theta^3$ asymptotic term. The integrals for the core and scattering terms are computed numerically, but we derive an analytic expression for the integral of the $1/\theta^3$ term, shown below.

Given a fit amplitude α , the Hankel transform for the $1/\theta^3$ term may be written as

$$\mathcal{F}_{1/\theta^3}(k) = \alpha \int_0^\infty \theta \left(\frac{1}{\theta^3}\right) J_0(\theta \ell) d\theta = \alpha \int_0^\infty \frac{J_0(\theta \ell)}{\theta^2} d\theta. \quad (\text{B16})$$

The analytic expression we use for this integral is

$$\int \frac{J_0(\theta \ell)}{\theta^2} d\theta = \ell \left[J_1(\theta \ell) - J_0(\theta \ell) \left(\frac{\theta^2 \ell^2 + 1}{\theta \ell} \right) - \frac{\pi \theta \ell}{2} \left(H_0(\theta \ell) J_1(\theta \ell) - H_1(\theta \ell) J_0(\theta \ell) \right) \right], \quad (\text{B17})$$

where $H_n(x)$ is the Struve function.

C. BEAM SOLID ANGLES FOR DIFFERENT EFFECTIVE FREQUENCIES

To indicate the effect of the passbands on the beams, we tabulate the solid angles for select effective frequencies. The effective frequencies of the band centers are currently uncertain to approximately 2.4 GHz. However, since this is in part due to systematic errors in the measurements, the relative uncertainties are smaller.

Similar to the correction we make to the main beams for use with the CMB, in each case we take the beam to be $B'(\ell) = B(\ell\nu_{\text{RJ}}/\nu_{\text{S}})$, where ν_{RJ} is the effective frequency for radiation with a Rayleigh-Jeans spectrum and ν_{S} is the effective frequency for the source of interest (either CMB, synchrotron emission, dust, or the thermal Sunyaev-Zel'dovich (tSZ) effect).

As described in §4.4, for the beam analysis for DR4, the effective frequencies from Thornton et al. (2016) were used for the RJ-to-CMB beam spectral correction. Subsequently, the effective frequencies were re-computed for the foreground modeling for DR4, using improved passband data and upgraded code, as detailed in Appendix D of Choi et al. (2020). These updated frequencies are shown here in Table 5 and the corresponding beam solid angles are shown in the following tables. Considering the uncertainties on the passbands, the frequencies from Thornton et al. (2016) and Choi et al. (2020) are consistent. In addition, given that the uncertainty on the beams is subdominant in the power spectrum analysis, which of these effective frequencies one uses for the beam spectral correction does not have a significant effect on the results. Still, the solid angles for the CMB in Table 6 below are slightly different from those in Table 4.

For any particular season/region/detector array/frequency, the uncertainty on the solid angles is 2–5%. This includes both the instantaneous beam uncertainty and the uncertainty due to the region-dependent jitter correction. As is apparent here, the derived solid angle and passband are intimately connected.

In addition, the beam and the passband are coupled, an effect which we considered for the first time in detail in Madhavacheril et al. (2020), and described in Appendix A of said paper. In short, the beam shape evolves as a function of frequency across the passband, and so our customary assumption of separability of these two components does not hold to high precision. This beam-bandpass coupling was modeled as part of a systematic check of the DR4 results in Choi et al. (2020) and was found to not have a significant effect on the inferred cosmological parameters.

TABLE 5
EFFECTIVE FREQUENCIES [GHz].

Array	Band	RJ	CMB	tSZ	dust	sync
PA1	150 GHz	150.8	149.6	150.0	151.2	146.9
PA2	150 GHz	151.2	149.9	150.4	151.6	147.3
PA3	150 GHz	148.4	147.6	147.9	148.7	145.8
PA3	98 GHz	98.7	97.9	98.4	98.8	95.5

TABLE 6
EFFECTIVE BEAM SOLID ANGLES (NSR) FOR THE CMB.

Array	Band	Season	Deep1	Deep5	Deep6	Deep56	Deep8	BOSS	AdvACT
PA1	150 GHz	S13	213	209	210	-	-	-	-
		S14	-	-	-	204	-	-	-
		S15	-	-	-	206	199	205	-
PA2	150 GHz	S14	-	-	-	190	-	-	-
		S15	-	-	-	196	192	193	-
		S16	-	-	-	-	-	-	194
PA3	150 GHz	S15	-	-	-	274	268	285	-
		S16	-	-	-	-	-	-	242
PA3	98 GHz	S15	-	-	-	546	541	556	-
		S16	-	-	-	-	-	-	524

TABLE 7
EFFECTIVE BEAM SOLID ANGLES (NSR) FOR SYNCHROTRON EMISSION.

Array	Band	Season	Deep1	Deep5	Deep6	Deep56	Deep8	BOSS	AdvACT
PA1	150 GHz	S13	221	217	218	-	-	-	-
		S14	-	-	-	211	-	-	-
		S15	-	-	-	213	207	212	-
PA2	150 GHz	S14	-	-	-	196	-	-	-
		S15	-	-	-	203	198	200	-
		S16	-	-	-	-	-	-	201
PA3	150 GHz	S15	-	-	-	281	275	292	-
		S16	-	-	-	-	-	-	248
PA3	98 GHz	S15	-	-	-	575	569	585	-
		S16	-	-	-	-	-	-	552

TABLE 8
EFFECTIVE BEAM SOLID ANGLES (NSR) FOR DUSTY SOURCES.

Array	Band	Season	Deep1	Deep5	Deep6	Deep56	Deep8	BOSS	AdvACT
PA1	150 GHz	S13	208	205	205	-	-	-	-
		S14	-	-	-	199	-	-	-
		S15	-	-	-	201	195	200	-
PA2	150 GHz	S14	-	-	-	186	-	-	-
		S15	-	-	-	192	187	189	-
		S16	-	-	-	-	-	-	190
PA3	150 GHz	S15	-	-	-	270	264	281	-
		S16	-	-	-	-	-	-	239
PA3	98 GHz	S15	-	-	-	536	531	546	-
		S16	-	-	-	-	-	-	515

TABLE 9
EFFECTIVE BEAM SOLID ANGLES (NSR) FOR THE TSZ EFFECT.

Array	Band	Season	Deep1	Deep5	Deep6	Deep56	Deep8	BOSS	AdvACT
PA1	150 GHz	S13	212	208	209	-	-	-	-
		S14	-	-	-	202	-	-	-
		S15	-	-	-	204	198	204	-
PA2	150 GHz	S14	-	-	-	189	-	-	-
		S15	-	-	-	195	190	192	-
		S16	-	-	-	-	-	-	193
PA3	150 GHz	S15	-	-	-	273	267	284	-
		S16	-	-	-	-	-	-	241
PA3	98 GHz	S15	-	-	-	541	536	551	-
		S16	-	-	-	-	-	-	520

D. TRANSFORMING FROM $\{Q_r, U_r\}$ TO $\{E, B\}$

Instead of estimating the polarized ℓ -space beams directly from non-local E and B transform maps, we opt to use a locally defined map-space polarization basis for this purpose. As we show below, the fields Q_r and U_r turn out to be a convenient choice; in the flat-sky limit, they are defined in terms of local linear combinations of the usual Q and U maps as follows:

$$Q_r(\boldsymbol{\theta}) = Q(\boldsymbol{\theta}) \cos 2\phi_\theta + U(\boldsymbol{\theta}) \sin 2\phi_\theta \quad (\text{D1})$$

$$U_r(\boldsymbol{\theta}) = U(\boldsymbol{\theta}) \cos 2\phi_\theta - Q(\boldsymbol{\theta}) \sin 2\phi_\theta \quad (\text{D2})$$

where $\boldsymbol{\theta} \equiv (\theta, \phi_\theta)$ are standard polar coordinates with the beam centroid as their origin and ϕ_θ increasing clockwise from the positive y -axis (assuming that one uses the convention in which $+x$ points to the right and $+y$ points upward). Conversely, we may also define these fields in terms of their local contributions to both Q and U in the same coordinate system:

$$Q(\boldsymbol{\theta}) = Q_r(\boldsymbol{\theta}) \cos 2\phi_\theta - U_r(\boldsymbol{\theta}) \sin 2\phi_\theta \quad (\text{D3})$$

$$U(\boldsymbol{\theta}) = U_r(\boldsymbol{\theta}) \cos 2\phi_\theta + Q_r(\boldsymbol{\theta}) \sin 2\phi_\theta . \quad (\text{D4})$$

Since we are ultimately interested in how leakage manifests itself in the usual angular power spectra, we need to translate any polarized beam models of Q_r and U_r to an ℓ -space representation of E and B . As it turns out, there exists a simple relation between the azimuthally averaged versions of these components, which we derive here in the flat-sky limit. We begin with the Fourier-space expressions for E and B :

$$E(\boldsymbol{\ell}) = \hat{Q}(\boldsymbol{\ell}) \cos 2\phi_\ell + \hat{U}(\boldsymbol{\ell}) \sin 2\phi_\ell \quad (\text{D5})$$

$$B(\boldsymbol{\ell}) = \hat{U}(\boldsymbol{\ell}) \cos 2\phi_\ell - \hat{Q}(\boldsymbol{\ell}) \sin 2\phi_\ell \quad (\text{D6})$$

where $\boldsymbol{\ell} \equiv (\ell, \phi_\ell)$ is the Fourier conjugate of $\boldsymbol{\theta}$, and $\{\hat{Q}, \hat{U}\}$ are just standard Fourier transforms of $\{Q, U\}$:

$$\hat{Q}(\boldsymbol{\ell}) = \int Q(\boldsymbol{\theta}) e^{i\boldsymbol{\ell} \cdot \boldsymbol{\theta}} d\boldsymbol{\theta} = \int Q(\theta, \phi_\theta) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta \quad (\text{D7})$$

$$\hat{U}(\boldsymbol{\ell}) = \int U(\boldsymbol{\theta}) e^{i\boldsymbol{\ell} \cdot \boldsymbol{\theta}} d\boldsymbol{\theta} = \int U(\theta, \phi_\theta) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta . \quad (\text{D8})$$

Taking the azimuthal average of Equations **D5** and **D6**, we get the one-dimensional transforms \tilde{E} and \tilde{B} :

$$\tilde{E}(\ell) = \frac{1}{2\pi} \int \hat{Q}(\boldsymbol{\ell}) \cos 2\phi_\ell d\phi_\ell + \frac{1}{2\pi} \int \hat{U}(\boldsymbol{\ell}) \sin 2\phi_\ell d\phi_\ell \quad (\text{D9})$$

$$\tilde{B}(\ell) = \frac{1}{2\pi} \int \hat{U}(\boldsymbol{\ell}) \cos 2\phi_\ell d\phi_\ell - \frac{1}{2\pi} \int \hat{Q}(\boldsymbol{\ell}) \sin 2\phi_\ell d\phi_\ell . \quad (\text{D10})$$

The expression for \tilde{E} in Equation **D9** may be rewritten in terms of map-space Q and U with the help of Equations **D7** and **D8**:

$$\tilde{E}(\ell) = \frac{1}{2\pi} \int Q(\theta, \phi_\theta) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta \cos 2\phi_\ell d\phi_\ell + \frac{1}{2\pi} \int U(\theta, \phi_\theta) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta \sin 2\phi_\ell d\phi_\ell . \quad (\text{D11})$$

Substituting for $Q(\boldsymbol{\theta})$ and $U(\boldsymbol{\theta})$ in Equation **D11** using Equations **D3** and **D4**, we obtain a relation between \tilde{E} and $\{Q_r, U_r\}$:

$$\begin{aligned} \tilde{E}(\ell) &= \frac{1}{2\pi} \int (Q_r(\theta, \phi_\theta) \cos 2\phi_\theta - U_r(\theta, \phi_\theta) \sin 2\phi_\theta) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta \cos 2\phi_\ell d\phi_\ell \\ &\quad + \frac{1}{2\pi} \int (U_r(\theta, \phi_\theta) \cos 2\phi_\theta + Q_r(\theta, \phi_\theta) \sin 2\phi_\theta) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta \sin 2\phi_\ell d\phi_\ell . \end{aligned} \quad (\text{D12})$$

Grouping together the terms with $Q_r(\boldsymbol{\theta})$ and $U_r(\boldsymbol{\theta})$, the equation above becomes:

$$\begin{aligned} \tilde{E}(\ell) &= \frac{1}{2\pi} \int Q_r(\theta, \phi_\theta) (\cos 2\phi_\theta \cos 2\phi_\ell + \sin 2\phi_\theta \sin 2\phi_\ell) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta d\phi_\ell \\ &\quad + \frac{1}{2\pi} \int U_r(\theta, \phi_\theta) (\cos 2\phi_\theta \sin 2\phi_\ell - \sin 2\phi_\theta \cos 2\phi_\ell) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta d\phi_\ell . \end{aligned} \quad (\text{D13})$$

Then, making use of a simple trigonometric identity, Equation D13 may be written as:

$$\begin{aligned}\tilde{E}(\ell) &= \frac{1}{2\pi} \int Q_r(\theta, \phi_\theta) \cos 2(\phi_\theta - \phi_\ell) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta d\phi_\ell \\ &\quad - \frac{1}{2\pi} \int U_r(\theta, \phi_\theta) \sin 2(\phi_\theta - \phi_\ell) e^{i\ell\theta \cos(\phi_\theta - \phi_\ell)} \theta d\theta d\phi_\theta d\phi_\ell.\end{aligned}\tag{D14}$$

And making the substitution $\phi_\rho \equiv \phi_\theta - \phi_\ell$, Equation D14 may be expressed as:

$$\begin{aligned}\tilde{E}(\ell) &= \frac{1}{2\pi} \int Q_r(\theta, \phi_\theta) \cos 2\phi_\rho e^{i\ell\theta \cos \phi_\rho} \theta d\theta d\phi_\theta d\phi_\rho \\ &\quad - \frac{1}{2\pi} \int U_r(\theta, \phi_\theta) \sin 2\phi_\rho e^{i\ell\theta \cos \phi_\rho} \theta d\theta d\phi_\theta d\phi_\rho.\end{aligned}\tag{D15}$$

We are now able to write each of the two terms in the expression for \tilde{E} as three separate integrals:

$$\begin{aligned}\tilde{E}(\ell) &= \int \left(\frac{1}{2\pi} \int Q_r(\theta, \phi_\theta) d\phi_\theta \right) \left(\int \cos 2\phi_\rho e^{i\ell\theta \cos \phi_\rho} d\phi_\rho \right) \theta d\theta \\ &\quad - \int \left(\frac{1}{2\pi} \int U_r(\theta, \phi_\theta) d\phi_\theta \right) \left(\int \sin 2\phi_\rho e^{i\ell\theta \cos \phi_\rho} d\phi_\rho \right) \theta d\theta.\end{aligned}\tag{D16}$$

Note that the integrals of $Q_r(\theta)$ and $U_r(\theta)$ over ϕ_θ — the first set of parentheses — are simply the azimuthal averages \tilde{Q}_r and \tilde{U}_r , while the integrals over ϕ_ρ - the second set of parentheses - turn out to have simple analytic counterparts:

$$\int \cos 2\phi_\rho e^{i\ell\theta \cos \phi_\rho} d\phi_\rho = -2\pi J_2(\ell\theta)\tag{D17}$$

$$\int \sin 2\phi_\rho e^{i\ell\theta \cos \phi_\rho} d\phi_\rho = 0.\tag{D18}$$

So the azimuthally averaged ℓ -space E beam is simply the second-order Hankel transform of the azimuthally averaged map-space Q_r beam:

$$\tilde{E}(\ell) = -2\pi \int \tilde{Q}_r(\theta) J_2(\ell\theta) \theta d\theta.\tag{D19}$$

One can similarly show that the same relation exists between the azimuthally averaged ℓ -space B and map-space U_r beams:

$$\tilde{B}(\ell) = -2\pi \int \tilde{U}_r(\theta) J_2(\ell\theta) \theta d\theta.\tag{D20}$$

With Equations D19 and D20 in hand, we have a complete formalism for transforming the polarized beams using the $\{\tilde{Q}_r, \tilde{U}_r\}$ basis.

E. DR3 VS DR4 BEAMS

The beam transforms made publicly available for the DR3 and DR4 releases have been compared for each season, sky region, detector array, and frequency in common, as shown in Figure 21. Despite the changes in the analyses, the beam transforms are consistent.

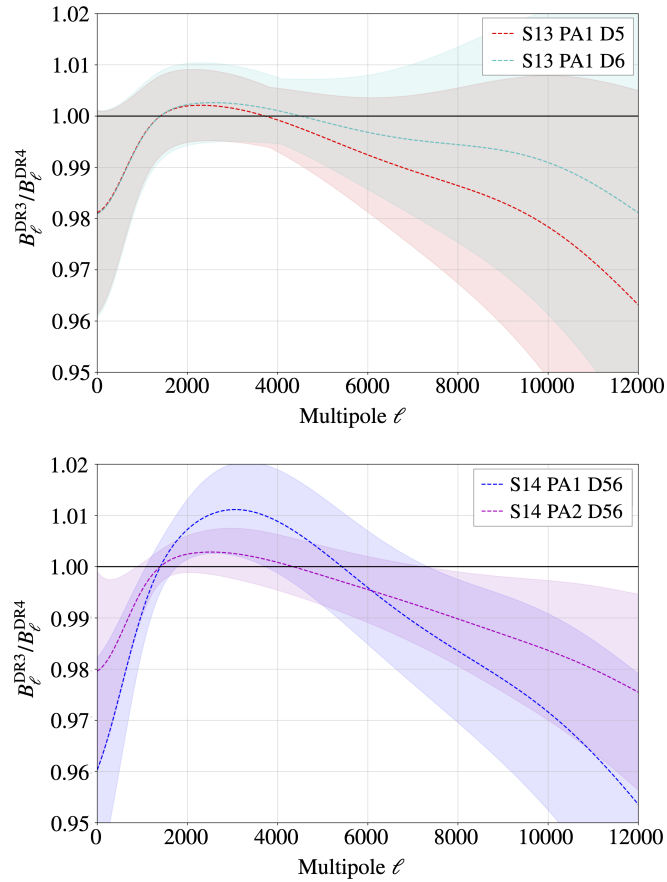


FIG. 21.— Ratio of the beam transforms for DR3 and DR4 at 150 GHz, for S13 (top) and S14 (bottom). The shaded bands indicate the 1σ uncertainty bounds, determined using the maximum uncertainty of the two transforms being compared. For these plots, the transforms have been normalized at $\ell = 1400$, which corresponds roughly to the effective calibration scale.