The Simons Observatory: Beam Characterization for the Small Aperture Telescopes

Nadia Dachlythra[1] , Adriaan J. Duivenvoorden[2,3] , Jon E. Gudmundsson[1,4] , Matthew Hasselfield[2] , Gabriele Coppi[5] ,
Alexandre E. Adler[1] , David Alonso[6] , Susanna Azzoni[6,7] , Grace E. Chesmore[8] , Giulio Fabbian[9,10] , Ken Ganga[11] ,
Remington G. Gerras[12] , Andrew H. Jaffe[13] , Bradley R. Johnson[14] , Brian Keating[15] , Reijo Keskitalo[16,17] ,
Theodore S. Kisner[16,17] , Nicoletta Krachmalnicoff[18,19,20] , Marius Lungu[21] , Frederick Matsuda[22] , Sigurd Naess[23] ,
Lyman Page[3] , Roberto Puddu[24] , Giuseppe Puglisi[25,26] , Sara M. Simon[27] , Grant Teply[15] , Tran Tsan[15] ,
Edward J. Wollack[28] , Kevin Wolz[6] , and Zhilei Xu[29]

[1] The Oskar Klein Centre, Department of Physics, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden
[2] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY, 10010, USA
[3] Joseph Henry Laboratories of Physics, Jadwin Hall, Princeton University, Princeton, NJ 08544, USA
[4] Science Institute, University of Iceland, 107 Reykjavik, Iceland
[5] Department of Physics, University of Milano—Bicocca, Piazza della Scienza, 3-20126 Milano (MI), Italy
[6] Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
[7] Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[8] Department of Physics, University of Chicago, 5720 South Ellis Avenue, Chicago, IL 60637, USA
[9] Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
[10] School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, UK
[11] Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France
[12] Department of Physics, University of Southern California, Los Angeles, CA 90089, USA
[13] Astrophysics Group and Imperial Centre for Inference and Cosmology, Department of Physics, Imperial College London, Blackett Laboratory, Prince Consort
Road, London SW7 2AZ, UK
[14] University of Virginia, Department of Astronomy, Charlottesville, VA 22904, USA
[15] Department of Physics, University of California, San Diego, La Jolla, CA 92093, USA
[16] Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[17] Department of Physics, University of California, Berkeley, CA 94720, USA
[18] International School for Advanced Studies (SISSA), Via Bonomea 265, I-34136 Trieste, Italy
[19] National Institute for Nuclear Physics (INFN)–Sezione di Trieste, Via Valerio 2, I-34127 Trieste, Italy
[20] Institute for Fundamental Physics of the Universe (IFPU), Via Beirut 2, I-34151 Grignano (TS), Italy
[21] New York, NY, USA
[22] Japan Aerospace Exploration Agency (JAXA), Institute of Space and Astronautical Science (ISAS), Sagamihara, Kanagawa 252-5210, Japan
[23] Institute of Theoretical Astrophysics, University of Oslo, Norway
[24] Instituto de Astrofísica and Centro de Astro-Ingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Macul, Santiago, Chile
[25] Università di Catania, Dipartimento di Fisica e Astronomia, Sezione Astrofisica, Via S. Sofia 78, I-95123 Catania, Italy
[26] INAF—Osservatorio Astrofisico di Catania, Via S. Sofia 78, I-95123 Catania, Italy
[27] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[28] NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA
[29] MIT Kavli Institute, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Abstract

We use time-domain simulations of Jupiter observations to test and develop a beam reconstruction pipeline for the Simons Observatory Small Aperture Telescopes. The method relies on a mapmaker that estimates and subtracts correlated atmospheric noise and a beam fitting code designed to compensate for the bias caused by the mapmaker. We test our reconstruction performance for four different frequency bands against various algorithmic parameters, atmospheric conditions, and input beams. We additionally show the reconstruction quality as a function of the number of available observations and investigate how different calibration strategies affect the beam uncertainty. For all of the cases considered, we find good agreement between the fitted results and the input beam model within an $\sim$1.5% error for a multipole range $\ell = 30$–700 and an $\sim$0.5% error for a multipole range $\ell = 50$–200. We conclude by using a harmonic-domain component separation algorithm to verify that the beam reconstruction errors and biases observed in our analysis do not significantly bias the Simons Observatory r-measurement

1. Introduction

The temperature anisotropy of the cosmic microwave background (CMB) has been mapped across a wide range of angular scales (see, e.g., Bennett et al. 2013; Planck Collaboration I et al. 2020). Information in the polarization anisotropies, which are significantly weaker, has yet to be characterized as extensively. Continued measurements of the CMB polarization will help break the degeneracy between various cosmological parameters and provide an additional probe into the cosmic inflation paradigm. For the latter case, the community is focusing on measuring the power of the parity-odd polarization component, the so-called B-mode polarization, on degree scales and larger

that could be directly sourced by a primordial background of stochastic gravitational waves, a key prediction of some inflationary scenarios (see, e.g., Komatsu 2022). It is common practice to quantify the amplitude of the primordial B-mode polarization in terms of the tensor-to-scalar ratio, r.

The Simons Observatory (SO) Small Aperture Telescopes (SATs) aim to constrain the tensor-to-scalar ratio with unprecedented sensitivity, targeting a statistical error of $\sigma(r) = 0.003$ (Simons Observatory Collboration et al. 2019) or better. In order to do so, a collection of 42 cm aperture SATs will observe the CMB temperature and polarization from a 5200 m altitude in the Atacama Desert in Chile. Observations will be done in six frequency bands to allow mitigation of Galactic foregrounds (see, e.g., Krachmalnicoff et al. 2016). The tensor-to-scalar ratio constraint is an ambitious goal that calls for a comprehensive understanding of our telescopes' performance.

Improper beam modeling can significantly bias the telescope's science goals. A small beam reconstruction error between different frequency bands is important for the success of foreground component separation analyses. The B-modes from the polarized Galactic foregrounds are much stronger than the primordial signal we are seeking; thus, a slightly biased estimation of the amplitude of the foregrounds due to calibration mismatch can lead to an important bias on the tensor-to-scalar ratio. Furthermore, recovering the beam transfer function with a small error also facilitates the calibration against Planck data. This calibration will happen at intermediate angular scales where the smoothing effect of the beam is important. Biased estimates of the beams will again lead to relative biases between the frequency bands, which may significantly bias the inference of the primordial B-mode amplitude.

The main beam systematics represents only a small fraction of the long list of optical systematics that can impact cosmological analysis of data from small-aperture CMB telescopes. These include beam asymmetries of various types, beam sidelobes, polarization angle errors, internal reflection causing so-called ghosting, pointing errors, and half-wave plate (HWP)–related systematics, including spurious scan-synchronous effects. For efforts related to constraining the amplitude of primordial B-mode polarization, it is perceivable that all of the effects listed above could be nonnegligible. Many of these effects are discussed in the following publications: Shimon et al. (2008), Fraisse et al. (2013), Planck Collaboration VII et al. (2016), Salatino et al. (2018), BICEP2/Keck Array XI et al. (2019), Xu et al. (2020), Abitbol et al. (2021), and Duivenvoorden et al. (2021). The accurate determination of azimuthally averaged Stokes I beam profiles for SO SATs represents a necessary but not a sufficient requirement for accurate constraints of the amplitude of primordial B-mode polarization.

Beam calibration techniques for CMB telescopes have been investigated in a number of publications (see, e.g., Aikin et al. 2010; Keating et al. 2013; Hasselfield et al. 2013; Planck Collaboration VII et al. 2016; BICEP2/Keck Array XI et al. 2019; Lungu et al. 2022). This paper adds to the existing literature by investigating the observational requirements and capabilities for beam reconstruction for the SO SATs. Although optical design software can be used to predict the far-field beam response, the final beam model used for science analysis will rely heavily on planet observations that are made through

fluctuating atmosphere. It is therefore important to develop algorithms that accurately capture the details of such observations. This involves creating simulations that include realistic detector noise and atmospheric emission and using them to show how our beam reconstruction depends on observation time, the properties of atmospheric emission, and low-frequency thermal variations in our instrument.

The paper is structured as follows. Section 2 describes the SAT instrument design and physical optics models that will be used throughout the paper. Section 3 summarizes the simulation pipeline. The details of the analysis methods are described explicitly in Section 4, with key results summarized in Section 5. Section 6 offers conclusions and discussion.

2. Instrument Design and Beam Modeling

The SATs use a three-lens cryogenically cooled silicon refractor design (Matsuda 2020). The optics have a 42 cm diameter aperture and support a wide field of view (35°; Galitzki 2018). The cryomechanical and optical design is described in Ali et al. (2020). Each SAT can support up to approximately 10,000 dichroic detectors occupying a total of seven hexagonal, 150 mm diameter silicon wafers forming a focal plane (also hexagonal) with a radius of approximately 17.5 cm (see Figure 1 of Galitzki 2018). The detectors are cooled to 100 mK, while the lenses and aperture stop are cooled to 1 K. The dichroic detectors operate at two low-frequency, two mid-frequency (MF), and two ultrahigh-frequency (UHF) bands centered near 27 and 39, 93 and 145, and 225 and 280 GHz.

The SATs are equipped with a cryogenically cooled HWP mounted skyward of the optics. The spinning HWP (at 2 Hz) modulates the linearly polarized component of the sky in a controlled fashion. Any unpolarized signal from the sky and atmosphere is left unmodulated, which suppresses temperature-to-polarization (T-to-P) leakage and allows for a clean measurement of the polarized sky signal.

More details about the HWP design and related studies for SO can be found in Hill et al. (2018) and Salatino et al. (2018). The telescopes are externally baffled to suppress signal from the ground and nearby mountains. Specifically, each SAT telescope has a freestanding ground shield, a nominally reflective comoving shield, and a nominally absorptive forebaffle. Diffraction caused by baffling elements can potentially create polarized beam sidelobes that couple to both the ground and the galaxy; modeling of the effect for a shielded refractor has been investigated in Adler & Gudmundsson (2020).

The beam models for the SO telescopes are generated using Ticra Tools[30] (formerly GRASP), proprietary software based on physical optics and the physical theory of diffraction. With Ticra Tools, we simulate various optical components, such as lenses, antennas, feed horns, and stops, allowing us to capture critical features of the SO SAT design. For this analysis, we simulate the 2D far-field co- and cross-polar beam maps for pixels at various locations on the focal plane. Figure 1 shows a representative telescope configuration as set up in Ticra Tools.

We simulate beam maps for four frequency bands centered on 93, 145, 225, and 280 GHz. We will be referring to those four frequency bands as mid-frequency 1 (MF1),

---

[30] TICRA, Landemaerket 29, Copenhagen, Denmark (https://www.ticra.com).
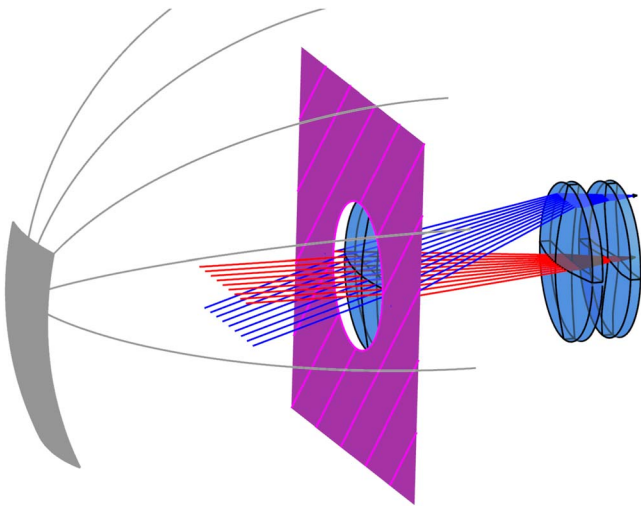
Figure 1. The three-lens SO SAT refracting telescope design that was implemented in Ticra Tools for the production of far-field beam maps for the SATs. In this setup, the light rays (in time-reversed simulations) travel from a center (red lines) and edge pixel (blue lines) of the focal plane through the three lenses (blue) and the aperture stop (purple surface) toward the far field (gray surface; not to scale), where the output beam is tabulated. The distance from the focal plane to the sky side of the primary lens is approximately 81 cm. The diameter of the three silicon lenses is about 45 cm.

mid-frequency 2 (MF2), ultrahigh-frequency 1 (UHF1), and ultrahigh-frequency 2 (UHF2). For each, we make band-integrated maps from five single-frequency simulations over a 20% bandwidth around the center frequency. For the nominal input models we use throughout the paper, we assume a top-hat spectral response function, although different weighting schemes can be implemented trivially. A class of potential input beam models for the SATs assuming nonuniform passbands and different types of frequency scaling are shown in Appendices A.1 and A.2 for reference.

Figure 2 shows the SAT beam profiles (top) and corresponding transfer functions (bottom) for the band-averaged simulations, assuming a pixel at the center (solid lines) and edge (dashed lines) of the focal plane, respectively. From the top panel of the figure, we see that the beam profiles are approximately Gaussian in the center with a sidelobe at larger angles, where the beam power roughly drops as the inverse cube of the angle. The edge pixel is located 18 cm from the center of the focal plane, corresponding to a beam centroid that is shifted by ~17°.5 relative to the telescope boresight. The center and edge pixel beam models shown in this figure will be assigned to all detectors of the center and one of the edge wafers correspondingly when simulating time streams.

The simulated far-field maps correspond to the copolar component of the beam. The cross-polar component is small in amplitude and should be studied together with the instrument's HWP performance. As planets are mostly unpolarized, observed polarization would be the result of T-to-P leakage, which the HWP failed to prevent (see Section 5 of Lungu et al. 2022). This part of the analysis is left for future work.

The SAT beams are treated as azimuthally symmetric throughout the paper, a choice that is strongly motivated by the GRASP simulations. As we will be scanning the sky when both rising and setting, the cross-linking in temperature maps will, furthermore, symmetrize the beams. Any T-to-P leakage
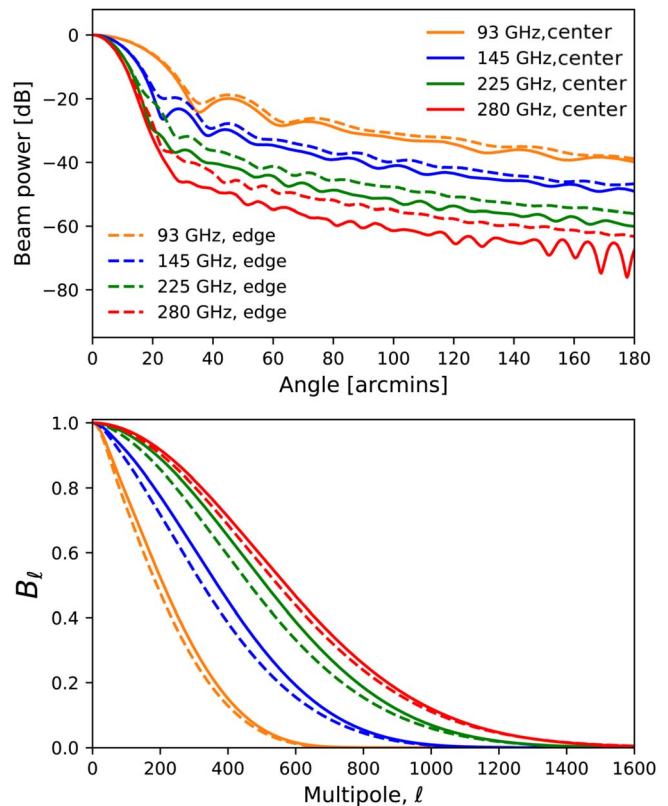
Figure 2. Top: beam profiles of the band-averaged input beam models for the four SO SAT frequency bands discussed in this paper. Solid and dashed lines show the cases for a detector placed on the center (0 cm) and the edge (18 cm) of the focal plane, respectively. The beam profiles are computed by radially binning the 2D band-averaged beam maps. Bottom: transfer functions for the beam models whose profiles are presented in the top panel.

Table 1
Best-fit Beam Size (FWHM), Ellipticity, and Solid Angle per Frequency Band for the Simulated SAT Beams

Frequency Band	FWHM (arcmin)	Ellipticity ϵ	Solid Angle $(10^{-6}$ sr$)$
MF1 (93 GHz)	27.4	0.030	78.9
MF2 (145 GHz)	17.6	0.036	30.5
UHF1 (225 GHz)	13.5	0.046	17.3
UHF2 (280 GHz)	12.1	0.045	13.6

Note. Assuming a pixel placed at the focal plane center. The forward gain derived from the total beam solid angle is 52.0, 56.2, 58.6 and 59.7 dBi for MF1, MF2, UHF1, and UHF2, respectively.

caused by remaining beam asymmetry is expected to be suppressed by the spinning HWP (Salatino et al. 2018).

Table 1 provides an overview of the simulated beams in terms of the full width at half-maximum (FWHM), solid angle, and best-fit value for the beam ellipticity, defined as

$$\epsilon = \frac{\theta_{\mathrm{maj}} - \theta_{\mathrm{min}}}{\theta_{\mathrm{maj}} + \theta_{\mathrm{min}}}, \qquad (1)$$

where θ_{maj} and θ_{min} are the FWHM of the beam's major and minor axis, respectively. The values in Table 1 are determined by fitting 2D elliptical Gaussians to the beam maps and apply to a pixel placed on the center of the focal plane. Our estimates suggest that the beam FWHM of a pixel located at the edge of

the focal plane will differ by about 1%–2% from its center pixel value, while the beam ellipticity may change by a factor of ~50% from the center to the edge of the focal plane. The results presented in this paper, however, are shown to be largely insensitive to the predicted variation in beam ellipticity across the focal plane (see Section 5.2).

3. Simulation Pipeline

In this section, we discuss potential calibration sources and describe the software and scan strategy employed to simulate the time domain data from observations of a bright point source, given the beam model discussed in the previous section.

3.1. Candidate Sources

Due to beam dilution, only a handful of natural point sources exist that are bright enough to calibrate the SAT beams. Of these, Jupiter is the brightest (Weiland et al. 2011; Planck Collaboration LII et al. 2017) and most suitable for SAT calibrations from the Atacama Desert. We focus on the characterization of the instrument's response to an unpolarized source as a baseline. The polarization response, which additionally requires a measurement of the instrument's polarization angle and cross-polar beam components, is left for future work.

Artificial calibration sources have the potential to overcome the limitations of the astrophysical sources. At the moment, sources mounted on tall structures have been successfully used for beam calibration (BICEP2/Keck Array XI et al. 2019). For the future, balloons (Masi et al. 2006), drones (Dünner et al. 2021), and even satellites (Johnson et al. 2015) are being considered. The use of drones for calibration purposes is the subject of multiple active studies (Nati et al. 2017). Calibration sources mounted on drones can be tuned in brightness, frequency, and cadence in order to meet the calibration requirements of different instruments. Additionally, these sources can be equipped with a polarizing wire grid that facilitates the calibration of polarization intensity and angle (Dünner et al. 2020). This last aspect is particularly important, as there are very few polarized astrophysical sources that are bright enough to calibrate the polarization response of the SO SATs (the highest upper limit of the planet polarization fraction, p_{frac}, is assigned to Uranus and corresponds to $p_{frac} < 3.6\%$ at 100 GHz within 95% confidence limits; Planck Collaboration LII et al. 2017).

There is a trade-off between calibrating with astrophysical and man-made sources. In the case of drones, flight endurance, especially in the thin atmosphere at high altitudes, is one of the main obstacles to successful calibration campaigns for experiments such as the SO. After recent on-site testing, the maximum flight time for the drone has been established to be ~12 minutes (Coppi et al. 2022). For this reason, the nominal calibration strategy relies on the planets.

Figure 3 shows a comparison of the signal-to-noise ratio (S/N) we can achieve when observing some of the brightest planets and when observing a source mounted on a drone as a function of frequency. The S/N estimation relies on the source's power and takes into account the noise-equivalent power (NEP) and optical efficiency of the telescope while assuming the same integration time per pixel for all sources. When using the drone, we can expect a significantly higher
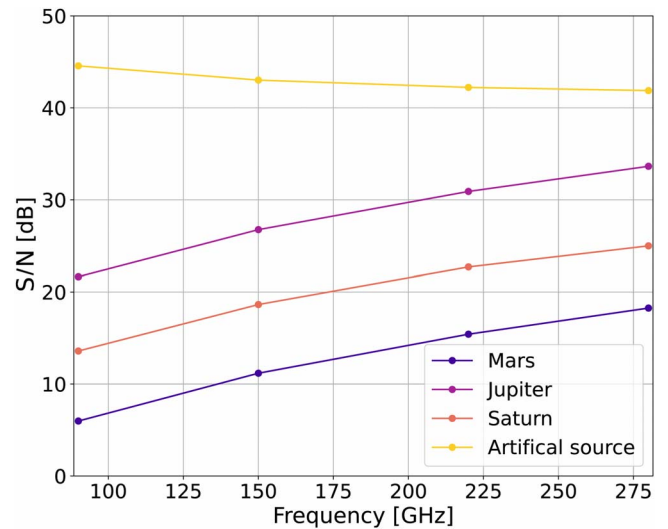
Figure 3. Comparison of the expected S/N when observing an artificial source mounted on a drone vs. planet observations for a single detector as a function of frequency. For all cases included in this figure, the noise has been considered to be in the range of 20–90 aW/$\sqrt{Hz}$, depending on the band, with a 1 s integration time and 20% bandwidth. The artificial source is assumed to be at a distance of 500 m from the telescope (with a power output of -18 dBm at all frequencies and an antenna gain of 6.5 dBi).

S/N, which scales more smoothly with frequency compared to the various planets' cases. Note that the planets' brightness values are calculated as a function of their average estimated distance from the Earth in 2023. The exact calculations leading to the S/N values of Figure 3 are described in Appendix B.

3.2. The TOAST and sotodlib Software

The simulated time-ordered data of the Jupiter observations are generated with the help of the Time-Ordered Astrophysics Scalable Tools (TOAST)[31] library, as well as the sotodlib[32] library, which interfaces with TOAST and provides experiment configuration files that are specific to SO.

The TOAST software was developed for simulating, gathering, and analyzing telescope time-ordered data. It is open-source software that is used in the framework of many current and next-generation CMB telescopes like LiteBIRD, Simons Array, SO, and CMB-S4. TOAST was used in a recent study of instrumental systematics for experiments aiming to observe the CMB polarization (Puglisi et al. 2021).

The software can generate instrumental noise, atmospheric noise, and scan-synchronous signals from ground pickup, as well as simulate the effect of an HWP. The code is end-to-end parallelized and optimized for low memory consumption, which facilitates its use with workload managers on large servers. TOAST allows one to simulate sky observations for different scan strategies of tunable parameters (scanning speed, observing time, and sky patch, among others) and implement various sky models.

The TOAST simulator module creates a focal plane configuration based on specified hardware parameters and samples the beam over a sky patch specified by the scheduler.

To simulate the atmosphere with TOAST, an additional file containing weather parameters is required. The corresponding module creates an atmospheric volume that moves with

---

[31] https://github.com/hpc4cmb/toast
[32] https://github.com/simonsobs/sotodlib
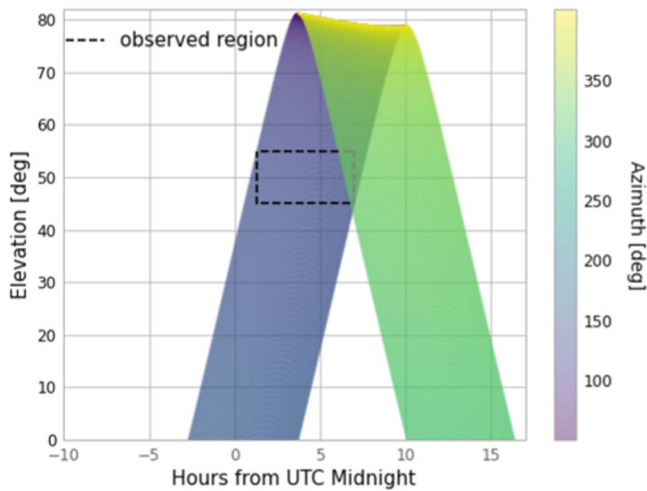
Figure 4. Jupiter's daily trajectory over the 3 months we are simulating (from 2021 June to August). Each of these curves shows the planet's elevation as a function of Coordinated Universal Time and is color-coded according to its azimuth value. The elevation of the observing region (indicated by the black dashed line) lies between 45° and 55° and corresponds to ∼1 hr of scanning.

constant wind speed, set by the user or randomly drawn, and observed by individual detectors on the focal plane. This part of the code needs a set of input parameters, for example, for the detector gain, the field of view, and the center and width of the dissipation and injection scales of the Kolmogorov turbulence, describing atmospheric fluctuations (Kolmogorov 1941; Errard et al. 2015). In this work, we limit ourselves to an atmospheric model based purely on water vapor, as experience proves it is the dominant disruption for CMB experiments (Morris et al. 2022). Other potential absorbers, such as clouds, ice crystals, and oxygen, are left for future work.

3.3. Scan Strategy and Noise Parameters

For the Jupiter observations, we simulate constant elevation scans (CESs) to avoid systematic effects arising from varying the telescope's elevation, and we set a maximum allowed observing time of an hour. Given an observing site, time, and target, TOAST simulates observing schedules that conform to observing constraints such as elevation and boresight distance to the Sun and Moon. It uses the PyEphem[33] software to predict the target's location with respect to the observatory. We run the scheduler with an allowed elevation range of [45°, 55°] and Sun/Moon avoidance radii set to 45°. The sampling rate and scanning speed are set to 200 Hz and $1°.5\,\mathrm{s}^{-1}$, respectively, at all times. The chosen maximum duration for a single CES facilitates tuning and calibrating the instrument; however, it takes Jupiter approximately 55 minutes to pass through the observing patch. After scheduling a single CES that follows these requirements, the scheduler moves on to the next day when Jupiter is available for observation. The simulated scans consider a single wafer (equipped with 860 detectors) each time instead of the full focal plane and cover an azimuth range of ∼20°. In practice, we will measure Jupiter with the HWP continuously spinning but we do not model it in the simulation. The signal amplitude from the sources is orders of magnitude greater than the polarized signal from the microwave sky; thus, we neglect this effect in the current simulations. Figure 4 shows

the daily trajectory of Jupiter over the full months of 2021 June, July, and August, which we choose to simulate in terms of the planet's elevation as a function of time. Each curve is color-coded according to Jupiter's azimuth value. The 2D interval defined by the black dashed line refers to the observed region.

It should be noted that the chosen azimuth and elevation ranges of the observing patch are not yet set in stone, and we expect them to evolve as we move closer to the telescope's deployment. This is because the choice of these parameters is primarily motivated by the source's availability during the observing period. The lower limit of the simulated elevation range is slightly lower than the most recent specifications, which set the nominal scanning elevation at 50°. As atmospheric loading worsens with decreasing elevation, our chosen strategy should be considered as a slightly pessimistic case, although we do not anticipate increasing the elevation range by 5° to improve our results significantly.

For the chosen observation period and scan strategy described above, we accumulated ∼50 hr of Jupiter simulations in total, as some days, the planet was not observable due to observing elevation and solar/lunar avoidance constraints. These simulations include atmospheric emission of both fixed and fluctuating weather parameters (see discussion in Section 5.3). For the scope of this project, we consider the atmosphere to be unpolarized even though it can intermittently carry a nonnegligible polarization fraction (Takakura et al. 2019). Furthermore, the simulations include realistic white and red detector noise expressed in noise-equivalent temperature. A detailed description of the noise model and parameters for the SATs can be found in Table 1 of Simons Observatory Collboration et al. (2019). For the planet temperature calculation, we rely on the thermodynamic temperatures listed in Table 4 of Planck Collaboration LII et al. (2017). The retrieved temperature values are then interpolated to the frequency range we wish to simulate and converted to T_{CMB} units, which express temperature in terms of the offset from the mean CMB temperature value.

4. Analysis Pipeline

We describe the mapmaker applied to the Jupiter simulations and the associated atmospheric noise mitigation techniques. We offer insights into the different parameters of the algorithm and summarize the method used to fit the radial beam profiles and transfer functions.

Both the mapmaking and beam fitting are based on methods developed for the Atacama Cosmology Telescope (ACT; Hasselfield et al. 2013; Lungu et al. 2022). The main new method development for this paper is the noise mitigation approach described in Section 4.1. Although the method is qualitatively similar to those used in the above-referenced papers, our new implementation allows for more fine-grained tuning of the noise subtraction compared to the ACT implementations. Additionally, the previous implementations have not been described in full detail in the literature, so we provide a detailed description here. The size of the telescope's field of view compared to the angular scale of the atmospheric fluctuations largely determines the effectiveness of the noise subtraction. A larger field of view will lower the effectiveness. We describe in detail how our implementation is tuned to take into account the large field of view of the SO SAT compared to ACT and demonstrate that the noise subtraction is still sufficiently effective.

---

[33] https://rhodesmill.org/pyephem/

4.1. Mapmaking and Low-level Processing

The simulation of planet observations starts with the generation of signal time streams by using the `TOAST` and `sotodlib` software. The time streams include white and correlated noise and are produced for all detectors of each of the seven SAT focal plane wafers, as described in Section 3.3. However, the results shown in this paper only concern the center and one of the edge wafers. We gather all of the time streams simulated this way and construct $10° × 10°$ maps around the planet.

We do not use a standard maximum-likelihood mapmaker for this analysis. While these are, in principle, unbiased and optimal and would therefore appear to be an ideal choice for measuring the instrument beam, in practice, they are only unbiased if the data precisely follow the fitted model. In reality, this is never the case. In this case, unmodeled gain and pointing fluctuations mean that the observed signal is time-dependent in a way that a static image of the sky cannot capture. The result of such model errors is a bias that is typically at the subpercent level but is nonlocal and spread out by roughly a noise correlation length. This bias is large enough to completely overwhelm the fainter wings of the beam profile (Næss 2019).

To avoid this bias, we use a specialized filter-and-bin mapmaker that uses our knowledge of the planet's position to build a filter that removes as much of the atmosphere as possible while leaving the planet's signal almost untouched (see Hasselfield et al. 2013; Choi et al. 2020). In particular, if we assume that the planet's signal is entirely contained inside a mask with radius θ_{mask} around its location, and that the noise is correlated with covariance matrix C_n, then we can use all of the data outside of the mask to make a prediction about the noise inside the mask and subtract it (Lungu et al. 2022):

$$d' = d - \underset{n}{\mathrm{argmax}}\, P(n|d_{\theta > \theta_{mask}}, C_n). \qquad (2)$$

Here d' and d are the raw and clean data vectors, respectively; n is the noise vector; and $d_{\theta > \theta_{mask}}$ are the data outside of the mask. To the extent that all of the signal is contained inside the mask, this subtraction will not introduce a bias. In practice, a small part of the signal will extend outside θ_{mask}, and there will be a trade-off between bias and noise subtraction (see discussion below).

For computational efficiency and to keep the implementation simple, we do not maximize P in Equation (2) but instead approximate it using the N_{modes} strongest principal components of a copy of d, where the area inside the mask has been filled in using polynomial interpolation. These principal components are then subtracted from the original d to form the cleaned d'. Effectively, we are using the detectors outside the mask at any given moment to predict what correlated noise the detectors inside the mask should be seeing and subtracting that. This approximation ignores the temporal correlations of the noise but seems to perform sufficiently well for our configuration.

After this cleaning, we assume that any remaining noise in d' is uncorrelated and can be mapped using a simple inverse-variance-weighted binned mapmaker. Note that the resulting map is only low-bias inside the masked region $\theta < \theta_{mask}$. Any data outside θ_{mask} are effectively high-pass-filtered due to the noise subtraction and thus heavily biased.

The effectiveness of this method depends strongly on the signal being compact (so one can use a small θ_{mask}) compared to the correlation length of the noise. It is therefore best suited
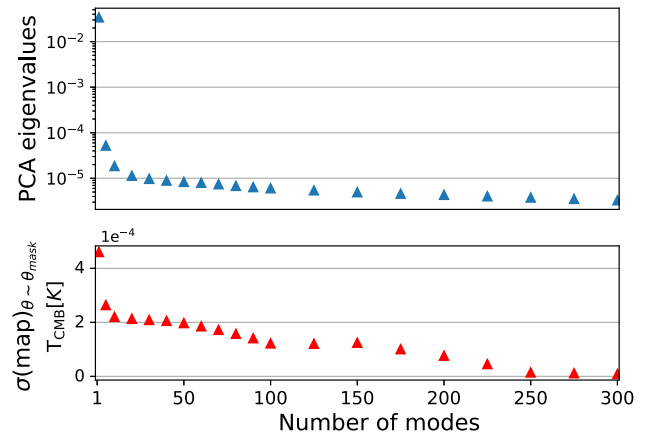

Figure 5. Top: the PCA eigenvalues of some of the 300 strongest correlated modes of the covariance matrix averaged over all detectors of the center wafer for a single 93 GHz Jupiter observation, as calculated from the mapmaker described in Section 4.1. Bottom: the noise amplitude of the binned data as a function of the number of correlated modes subtracted. The noise level is calculated as the standard deviation of all data at the outer 10% of the mask. Notice that the eigenvalues are shown in logarithmic scale, while the noise levels are in linear scale.

for high-resolution telescopes like ACT or the SO Large Aperture Telescope but still performs reasonably well for the SO SAT.

The mask radius, θ_{mask}, around the source and the number of modes, N_{modes}, can be tuned to optimize the performance of the algorithm. As the noise levels are estimated from the region that remains unmasked, a too-small θ_{mask} might result in subtracting beam power of substantial amplitude, while one might fail to properly capture the relevant noise modes with a too-large θ_{mask}. The top panel of Figure 5 shows the principal component analysis (PCA) eigenvalues of some of the 300 strongest modes for a single observation performed with the 93 GHz frequency band beam model and fixed precipitable water vapor (PWV). Even though the total number of estimated modes equals the number of simulated detectors (860 detectors per single-wafer simulation), we find this truncated sample representative enough to capture the rate of the eigenvalues' decreasing amplitude. Each data point in this plot corresponds to the average eigenvalues of all of the detectors of the center wafer of the focal plane. The large value of the first mode presented in this figure (∼2 orders of magnitude larger than the next mode) indicates how the atmospheric noise may be crudely approximated as a single correlated mode. The bottom panel of Figure 5 presents the noise amplitude of the same 93 GHz planet map estimated as the standard deviation of all of the data points included in the outer 10% of the mask as a function of the number of subtracted modes. The results illustrate an overall reduction of the noise amplitude with an increasing number of modes in an almost monotonic fashion. The noise variance is shown to be statistically compatible with zero after subtracting ∼300 modes.

The outer scale of atmospheric turbulence is observed to be significantly smaller than the 12° field of view of a single SAT wafer (Errard et al. 2015). Splitting the wafer into subsets of fewer detectors and estimating the correlated modes across these subsets instead would likely adequately capture the atmospheric noise model with a smaller number of modes than in the case of estimating the noise correlations from the full wafer. Relevant modifications to the mapmaking pipeline will be made, if necessary, when we start observations.
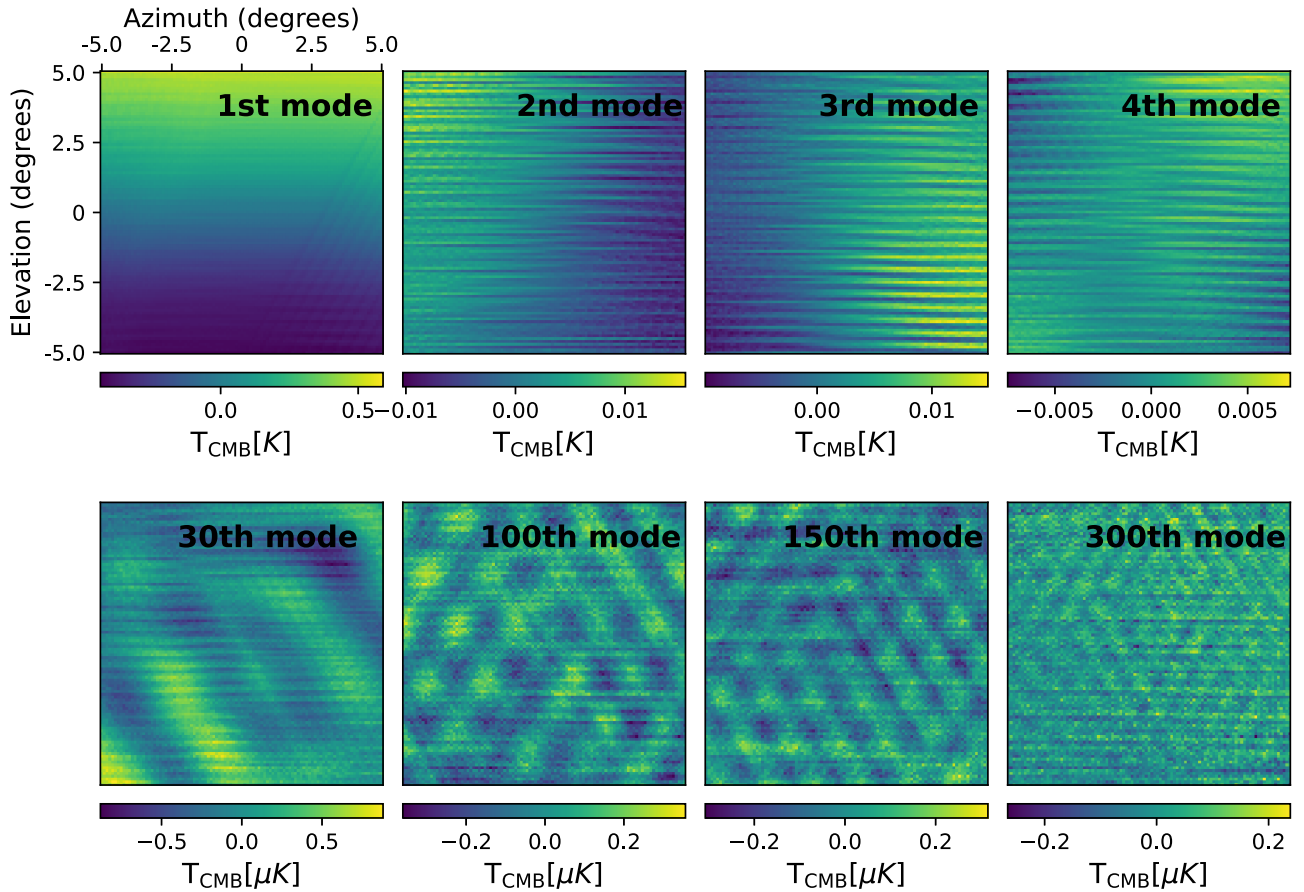
Figure 6. The binned 1st, 2nd, 3rd, 4th, 30th, 100th, 150th, and 300th strongest correlated modes of the covariance matrix of the center wafer detectors for a single 93 GHz observation, as calculated from the mapmaker described in Section 4.1. Note the change of units from K to μK between the top and bottom rows.

We set θ_{mask} to be equal to a radius, outside of which the beam power has fallen below $\sim$0.01% of its peak value, following the example of Lungu et al. (2022). Figure 6 shows some of the binned modes that were calculated from the PCA analysis for the same simulation. The first mode shows the atmospheric emission amplitude scaling with telescope boresight elevation, as expected, while the rest of the modes of the top row probe stripy patterns in slightly different directions and scales. The bottom row shows modes corresponding to detector correlations of significantly smaller amplitude. Subtracting these faintest modes might be excessive and could, in turn, end up negatively impacting the performance of the beam model reconstruction algorithm. Figure 5 suggests that the most suitable number of modes to subtract should be of the order of 10.

It is interesting to look at the impact of subtracting a different number of correlated modes directly on planet maps. Figure 7 shows a single, 93 GHz, $4° \times 4°$ planet map after subtracting the 1st, 2nd, 3rd, 4th, 30th, 100th, 150th, and 300th modes, following the reasoning of Figure 6. From this multipanel plot, we see the atmospheric striping starting to subside after the $\sim$fourth mode, while the contrast between the masked and unmasked region becomes more pronounced with increasing number of subtracted modes. As expected, the faintest eigenmodes of the covariance matrix will better capture the noise of the unmasked region, since the eigenmodes are estimated in this region. Consequently, subtracting these faint modes from the maps will lower the noise amplitude in the

unmasked region but not impact the masked region around the source as much.

4.2. Beam Profile Fitting

For the beam profile fitting of the planet observations, we closely follow the method developed for the ACT (Lungu et al. 2022). The beam fitting method is implemented in `beamlib`, a version of the code from the work of Lungu et al. (2022) that has been adjusted to SO hardware specifications. This section summarizes the main aspects of the method.

The fitting pipeline takes a set of input maps and trims them to a size that should refer to a region well contained inside θ_{mask}. The next step is to correct for the bias caused by the noise mode subtraction. As in the case of the ACT beams, we find this effect to be fairly consistent with a constant offset deviation of the beam wing from the $1/\theta^3$ function that the input SAT beams approximately follow (see Section 2). For the constant offset estimation, we use a relatively "flat" part of the beam profiles (where the oscillatory side-lobe pattern is not as pronounced), and the beam power has fallen below $\sim$0.1% of its peak value. The best-fit value for this offset is computed for each observation and then subtracted from each observation before averaging the beam profiles.

The core (main) beam is fitted, employing the basis functions from Lungu et al. (2022),

$$f_n(\theta\ell_{\mathrm{max}}) = \frac{J_{2n+1}(\theta\ell_{\mathrm{max}})}{\theta\ell_{\mathrm{max}}}, \qquad (3)$$
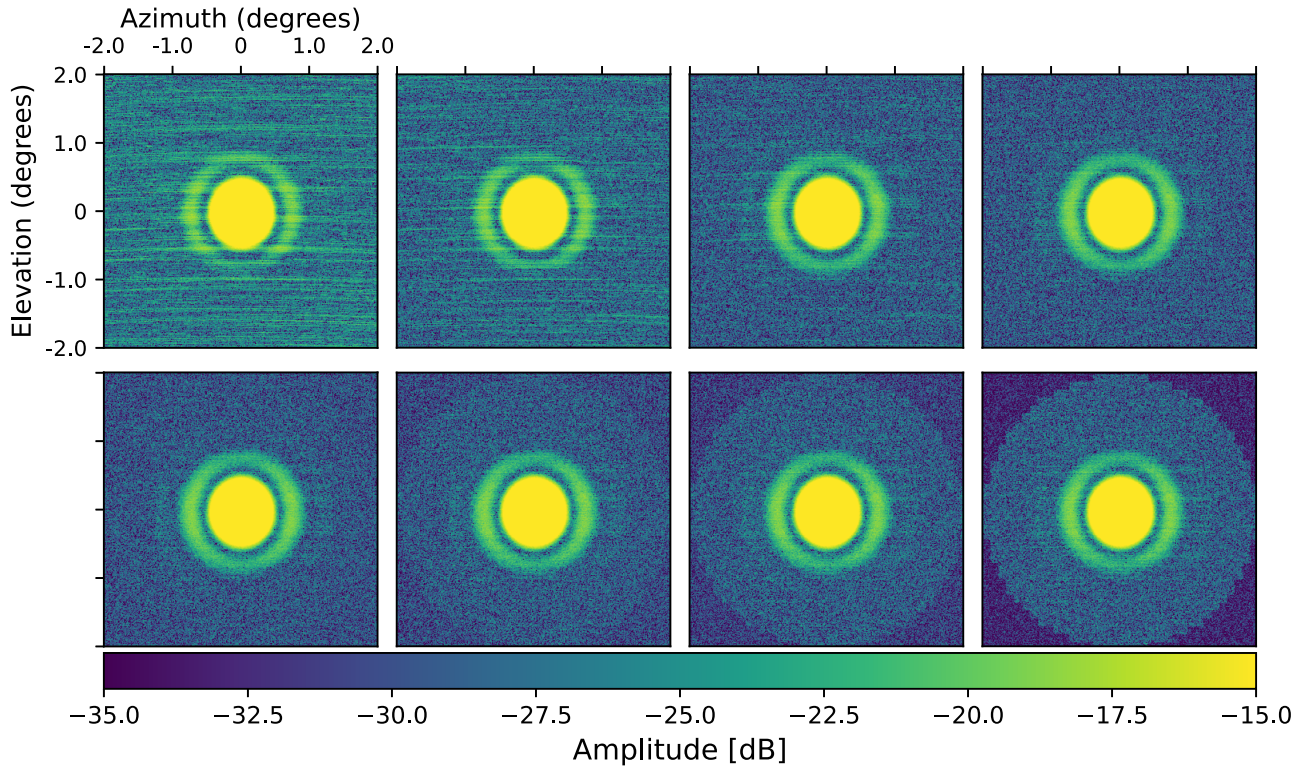
Figure 7. Peak-normalized beam maps constructed from the same 93 GHz Jupiter simulation after subtracting up to the 1st, 2nd, 3rd, 4th, 30th, 100th, 150th, and 300th strongest correlated modes shown in Figure 6. Note that now we only show $4° \times 4°$ patches around the source, as we want to take a closer look at the noise subtraction effect inside the masked region. The chosen color scale saturates the beam but better captures the noise mitigation progression.

where J_{2n+1} are Bessel functions of the first kind, n is a nonnegative integer, and ℓ_{max} is allowed to vary around some mean value defined by the beam resolution. The angle of the transition from the "core" to the "wing" region of the beam is specified by the user, yet it is allowed to vary slightly in the code in order to retrieve its optimal value. The reconstructed beam transfer function is calculated as the Legendre transform of the best-fit model,

$$b_\ell = \frac{2\pi}{\Omega_B} \int_{-1}^{-1} B(\theta) P_\ell(\cos\theta) d(\cos\theta), \quad (4)$$

where $P_\ell(\cos\theta)$ are the Legendre polynomials, Ω_B is the beam solid angle, and $B(\theta)$ is the best-fit radial beam profile comprised of the core and wing fit. Besides the best-fit harmonic transform, `beamlib` also calculates the eigenmodes of the beam fitting covariance matrix, which we will refer to as error modes throughout the paper.

A small number of available planet observations can be problematic when estimating the bin–bin covariance matrix of the binned radial profile. For that reason, the code employs a shrinking technique. The idea behind the approach is the following: the level of down-weighting of the off-diagonal components of the covariance matrix should depend on the number of available observations. This approximation is parameterized by the so-called shrinkage intensity, $\hat{\lambda}^*$ (see Equation (A5), Appendix A, Lungu et al. 2022). The shrinkage intensity is applied to a biased version of the covariance matrix, T, where we have set all the off-diagonal components to zero and the empirical, unbiased covariance matrix, S, to synthesize the "shrunk" covariance matrix, C, that we will use for the
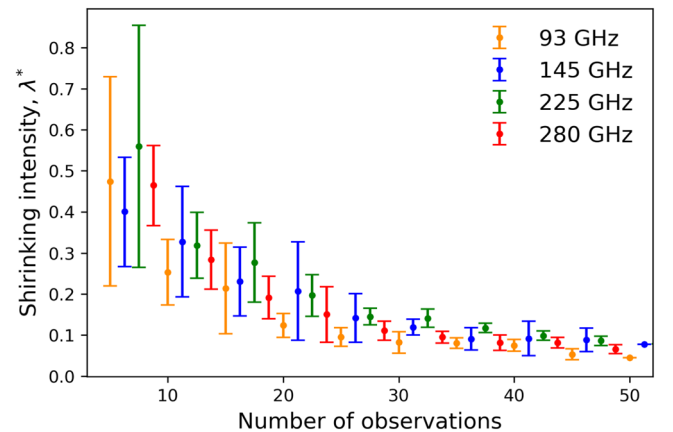
Figure 8. Shrinking estimator λ^* as defined in Lungu et al. (2022) as a function of the number of input Jupiter simulations for all frequency bands.

beam fitting as (Equation (A6), Appendix A, Lungu et al. 2022)

$$C = \hat{\lambda}^* T + (1 - \hat{\lambda}^*) S. \quad (5)$$

One can easily conclude that the larger the number of observations, the closer we can get to an unbiased covariance estimation. Figure 8 shows the value of λ^* as a function of the number of input observations to the beam fitting code for our frequency bands. The shrinkage intensity seems to converge to a value of ≈ 0.1 for a set of $N_{obs} = 50$ observations in all cases.

5. Results

We now present the reconstructed beam profiles and corresponding transfer functions for several different
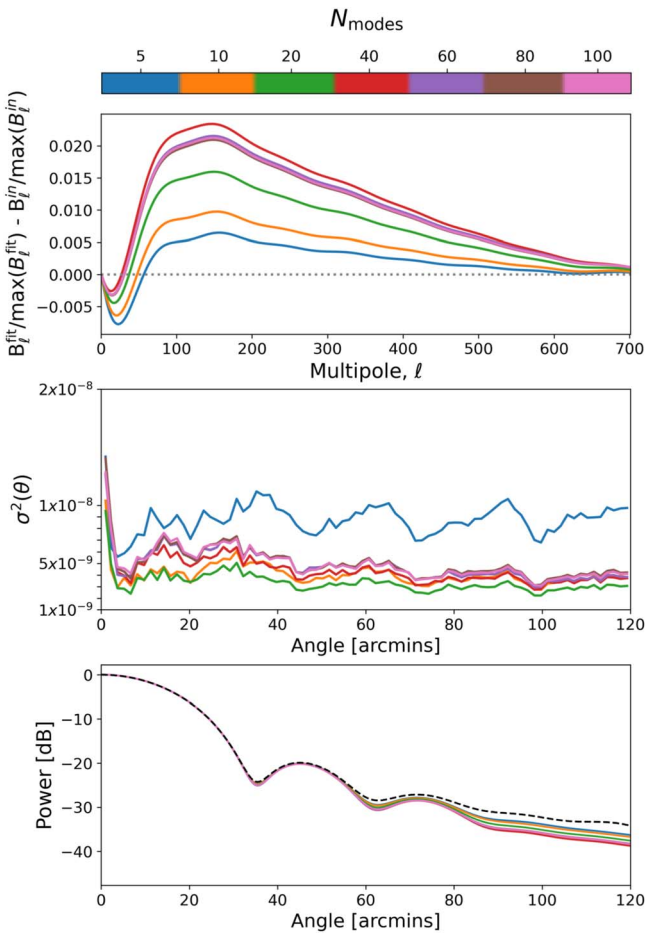
Figure 9. Top: the linear bias in multipole space between the normalized fitted and input beam profiles. Middle: the variance of the different planet maps (in logarithmic scale). Bottom: the best-fit beam profiles along with the input beam profile (black dashed line). All results are shown as a function of the number of correlated modes for the 93 GHz frequency band.

simulation parameters for the 93 GHz band. A subset of indicative results is shown for the rest of the frequency bands as well. We are particularly interested in isolating the factors that will most strongly impact the quality of our beam reconstruction.

5.1. Dependence on the Subtracted Correlated Noise Modes

The beam fitting performance is tightly linked to our ability to sufficiently suppress the atmospheric signal, which, in turn, depends on the number of correlated noise modes one removes from the data. For the simulation setup we employed, we find that the number of correlated modes that need to be removed typically lies between $N_{modes} = 5$ and 50 for the frequency bands we consider.

The choice of the number of subtracted modes relies on a combination of different criteria, as summarized in Figure 9. The top panel shows the linear bias in multipole space between normalized fitted and input beam transfer function as a function of the number of modes removed (varying colors) for simulations performed at the 93 GHz frequency band. The middle panel shows the corresponding real-space variance between planet maps for the different number of modes subtracted. Each curve shown in this panel is estimated by computing the per radial bin variance of the beam profiles of a

set of maps that have the same number of correlated modes removed. The bottom panel of the figure refers to the best-fit beam profiles of the same planet sets along with the input beam model.

From the middle panel, we see that the logarithmic variance between different observations consistently reduces with an increasing number of subtracted modes, but after some mode ($N_{modes} \gtrsim 20$), it starts increasing again. We expect the strongest noise modes to be quite similar between simulations assuming different dates, but fainter modes should reflect some degree of day-to-day atmospheric changes. Subtracting these fainter modes inevitably increases the variance between different maps to some extent. Such behavior can indicate that we should not remove any more modes to avoid the risk of also subtracting signal along with the noise.

The chosen number of modes, N_{modes}, should be the best combination of low bias and low variance, which is the case for $N_{modes} = 10$–20. Ideally, we would like the bias of the fitted beam transfer function to be fairly constant across the different multipole bins. We decide to use $N_{modes} = 10$ for the 93 GHz case.

Higher-frequency bands require more modes to be removed, since the atmospheric brightness scales with frequency. For 145, 225, and 280 GHz, we find that the number of modes that best satisfies our selection criteria is $N_{modes} = 10$, 30, and 40, respectively. Depending on the different simulation parameters, one might find a preferred (narrow) range of modes instead of a single global value. An alternative approach would be to subtract all of the correlated modes and estimate a transfer function for the bias caused by this process by running additional simulations. For this work, we aim to only mildly bias our data so that the nature of this bias is fairly predictable and, in turn, easily corrected.

5.2. Dependence on Detector Position

The fidelity of the beam reconstruction process depends on the input beam models. The input models depend, in turn, on the position of the detector on the focal plane. To assess the impact of the detector position on the fitting performance, we use beam models for a pixel on the center and the edge of the focal plane. In both cases, we bin the data into maps with the 10 most correlated modes removed after masking and gap-filling a region that extends to a radius of $\theta_{mask} = 2°.5$ around the source. The gap-filling is done with polynomial interpolation of the unmasked data over the masked region.

Figure 10 shows the reconstructed beam profiles from a set of 3 months of simulated Jupiter observations for the center (orange curve) and one of the edge wafers (green curve). The detectors of the center wafer share the center pixel beam model, while those of the edge wafer are assigned the edge pixel beam model (as discussed in Section 2). The selection of the exact edge wafer is not important, as the Ticra Tools setup is radially symmetric with respect to any of the edge pixels of the telescope's focal plane (see Figure 1). The atmospheric PWV was set to ~1 mm at all times, and we did not allow for wind speed variations. The results show the fitted center pixel beam model following the input model closely, up to ~four times the beam size (FWHM = 27′.4 at 93 GHz), while the edge pixel beam profile deviates slightly from the input toward the largest radial bins.

Figure 11 quantifies the above statement in terms of the bias on the beam transfer functions of the two reconstructed beam
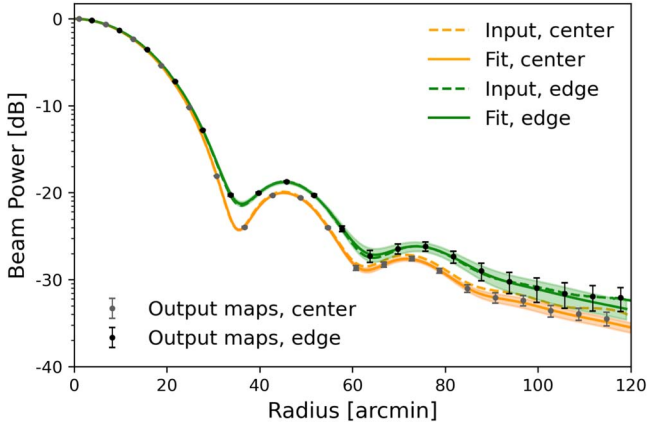
Figure 10. Best-fit beam profiles generated from 3 months of Jupiter simulations for the 93 GHz frequency band, performed with an input beam model for a center (orange curve) and an edge (green curve) pixel. The input beams (orange/green dashed lines) and data points (gray/black error bars) of the fitted models are also shown.
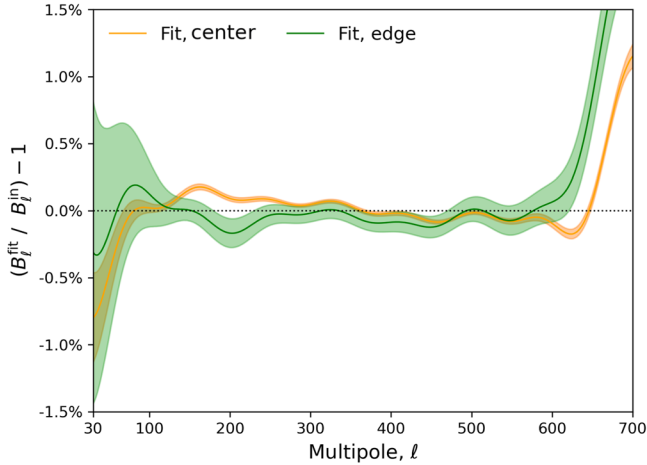

Figure 11. The beam transfer function bias, as compared to the input model, generated from 3 months of Jupiter simulations for the 93 GHz frequency band, performed with an input beam model for a center (orange curve) and an edge (green curve) pixel. The band around the solid lines represents the 1σ error envelope determined from the beam error modes.

profiles compared to the input beams. From the figure, we can see that, while the center pixel model reconstruction is better, the transfer function bias remains well under 1.5% for a multipole range of $\ell = 30$–700 for both cases.

5.3. Dependence on Weather Conditions

Different atmospheric conditions could affect the performance of the beam reconstruction algorithm. During the chosen observation period, we expect the weather conditions to vary to some extent. Generally speaking, there are a variety of parameters driving the atmospheric behavior. From these parameters, the amount of PWV has the strongest impact on the atmospheric emission (see, e.g., Dünner et al. 2012; Errard et al. 2015).

The impact of PWV on atmospheric transmission can be seen in Figure 1 of Errard et al. (2015). We define the transmission at some frequency, ν, as the ratio $\mathcal{T}(\nu) \equiv I(\nu)/I_0(\nu)$ of the radiation received by the detector, $I(\nu)$, and the radiation above the atmosphere, $I_0(\nu)$. A high PWV value implies a low transmission $\mathcal{T}(\nu)$, which can be defined as the
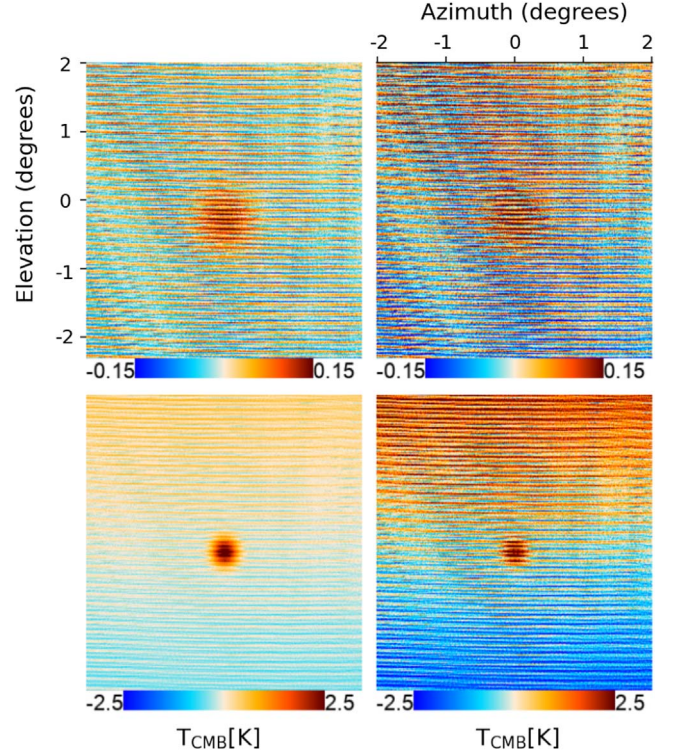
Figure 12. Binned data of a single, $\sim$1 hr, Jupiter observation for all detectors in the center wafer at 93 (top) and 280 (bottom) GHz that have the mean temperature subtracted and have no correlated noise modes removed. The left and right panels represent simulations that include atmospheric emission of PWV $= 0.5$ and 2.5 mm, respectively.

negative exponent of the air mass, $m(90° - el)$, times some standard value of the optical depth, τ_0, measured at the zenith (Errard et al. 2015),

$$\mathcal{T}(\nu) = e^{-m(90° - el)\tau_0}, \tag{6}$$

where el is the elevation. The air mass is, in turn, computed as a function of the zenith angle (see Equation (2) of the same paper). This relationship holds at high elevations if we model the atmosphere as a parallel planar slab; in this case, $m(90° - el) \approx 1/\sin(el)$. The atmospheric transmission contributes to the total loading, $\mathcal{E}(\nu)$, as follows (Errard et al. 2015):

$$\mathcal{E}(\nu) = [1 - \mathcal{T}(\nu)]B_\nu(T_{atm}). \tag{7}$$

In the above, $B_\nu(T_{atm})$ is the spectral radiance of a blackbody of temperature equal to the atmospheric temperature, T_{atm}.

Figure 12 shows the binned time-ordered data of a single Jupiter observation, including atmospheric emission of PWV $= 0.5$ mm (left column) and PWV $= 2.5$ mm (right column) at 93 (top row) and 280 (bottom row) GHz. The data are binned after subtracting the mean (atmospheric) temperature and have no correlated modes removed. The atmospheric intensity and therefore striping is shown to increase non-negligibly with PWV value, as expected. For the cases presented, the S/N at 0.5 (2.5) mm is estimated as S/N $= 35$ (30) for the 93 GHz band and S/N $= 140$ (60) for the 280 GHz band.

The atmospheric signal is strongly correlated between different detectors. The wind speed and direction impact the correlation length along with the outer scale of turbulence and

Table 2
PWV, Wind Speed, and Direction Assumed in the Various Simulation Cases
Described in Section 5.3

Case	PWV (mm)	South Wind Speed (m s^{-1})	West Wind Speed (m s^{-1})
i	1.17	-1.25	3.4
ii	1.17	-1.25 ± 1	3.4 ± 2.5
iii	2.5	-1.25	3.4
iv	1.17 ± 1	-1.25	3.4
v	1.17 ± 1	-1.25 ± 1	3.4 ± 2.5

scan strategy. For two detectors with beam centroids that lie parallel to the wind direction, we will observe maximum signal correlation (Equations (23)–(26) of Morris et al. 2022). To explore these effects, we run simulations for a center pixel in the 93 GHz frequency band and the different cases of atmospheric parameters described in Table 2.

The PWV value, wind speed, and direction are either fixed or allowed to fluctuate around a mean value following some distribution that is consistent with historical distributions according to MERRA-2 (Global Modeling & Assimilation Office (GMAO) 2015) for the Atacama observation site. These distributions are included in TOAST and specified per hour of the day and month. The mean and standard deviation values quoted in Table 2 are synthesized from the individual simulations of the full observing period we have chosen.

Notice that the estimated fixed mean PWV value strongly agrees with the one motivated by seasonal data of the ACT telescope ($\sim$1 mm), which is located at the SO observation site (see Figure 4 of Morris et al. 2022). The simulated PWV is uniformly distributed, and the surface temperature and pressure are kept constant at 270 K and 530 hPa, respectively. The wind speed values in Table 2 have a positive or negative sign in order to also incorporate the wind direction.

Figure 13 shows the uncertainty of the reconstructed beam transfer function with respect to the input beam model for cases (ii), (iii), (iv), and (v), as described in Table 2, in the form of blue, orange, green, and red error bars, respectively. The mean bias of the nominal case, (i), is demonstrated with a black dashed line surrounded by a gray shaded uncertainty band for reference. The error bars are shown per 40 multipoles for easier visualization and extend to a multipole number of $\ell_{max} = 700$. The beam fitting performance is rather stable across the different weather cases we have considered. However, the quality of the results slightly worsens with added atmospheric complexity, with the largest error bars of Figure 13 corresponding to the case where both PWV and wind speed are allowed to fluctuate (case (v)). A large PWV value implies an increased temperature of the atmospheric brightness. Since the latter also scales with frequency, the quality of the results depends on the ratio between the planet and atmospheric brightness at the frequency band of interest and how efficiently we can suppress the correlated noise. Nevertheless, we should highlight that, for all atmospheric parameters chosen, we were still able to recover the input beam with an uncertainty smaller than $\sim$1.5% in all cases for the multipole range $\ell = 30$–700 (for 93 GHz).

5.4. Dependence on the Number of Observations

The beam reconstruction algorithm depends on the number of available observations. To probe this, we present the
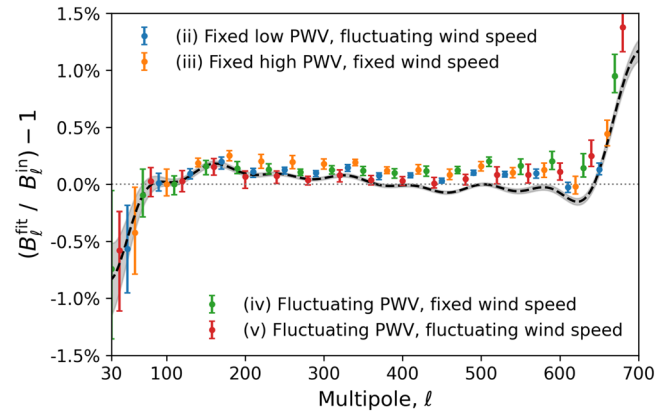

Figure 13. Beam transfer function bias from the input beam model for simulation cases (ii)–(v) described in Table 2. The simulations were performed with the 93 GHz frequency band beam model assuming a pixel at the center of the telescope's focal plane. Note that, for ease of visualization, the different error bars are slightly offset in the ℓ-direction. The black dashed case corresponds to the nominal case (i).
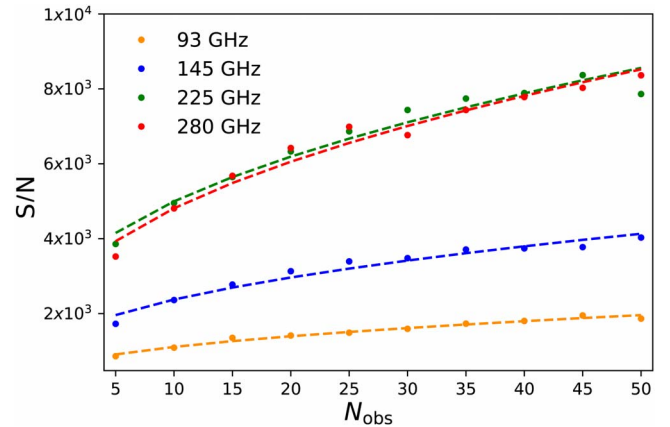
Figure 14. S/N as a function of the number of simulated $\sim$hour-long Jupiter observations for four frequency bands centered on 93, 145, 225, and 280 GHz. The circles represent S/N values estimated by determining the noise levels of the maps, and the dashed lines represent the best fits of the data points to an $A\sqrt{N_{obs}} + B$ model.

accumulated S/N as function of the number of Jupiter simulations for four different frequency bands centered on 93, 145, 225, and 280 GHz. The beam models in the simulations assume a detector placed at the center of the focal plane.

In our analysis, we face a trade-off between the overall accuracy of the reconstructed beam model and extending the model to larger angles. The S/N obviously decreases as we move away from the center of the beam. Therefore, our attempts at fitting beam models in the faint wings of the sidelobes can sometimes bias our overall results. Assessing the reconstruction noise as a function of the number of observations is essential for optimizing the planet observing strategy. Figure 14 shows the estimated S/N for all frequency bands when the number of available observations ranges from 5 to 50. The circles refer to S/N values estimated by determining the noise that remains in the planet maps, which is approximated as the standard deviation of the data in the outer 10% of the mask. The dashed lines are the fits of the S/N values to an underlying $A\sqrt{N_{obs}} + B$ model (for some constants A and B), which is the statistical behavior we would expect. Based on Figure 14, we decide that attempting to model all frequency bands down to
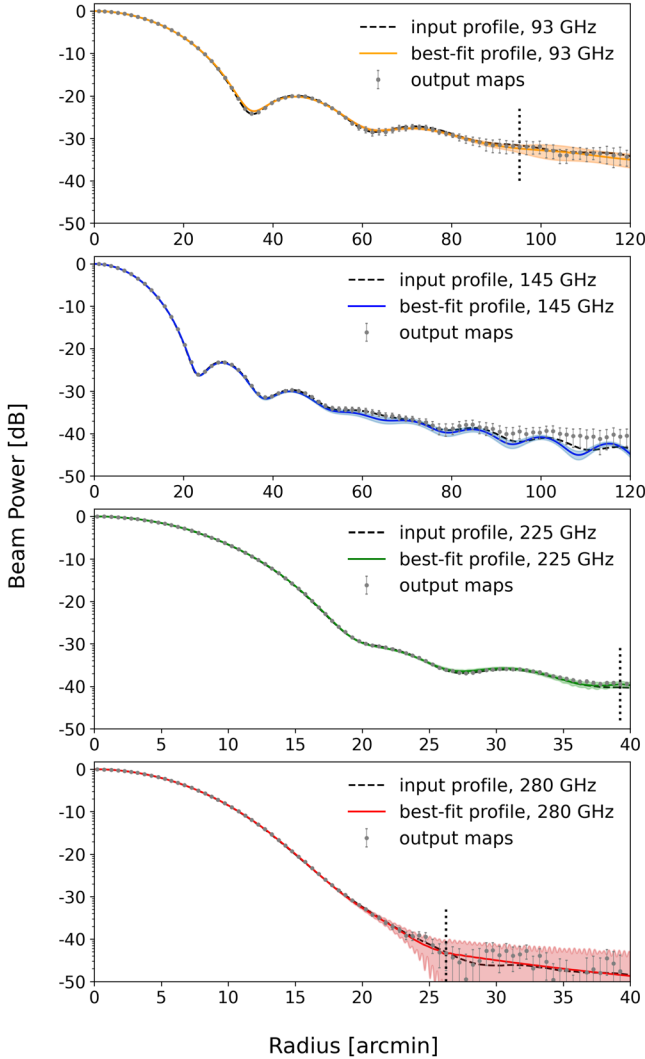
Figure 15. Best-fit beam profiles for the boresight pixel beam model at the 93, 145, 225, and 280 GHz frequency bands. Note that the panels showing the 93 and 145 GHz results have a different horizontal range than those showing results for 225 and 280 GHz. The vertical black line represents the transition between the core and wing fit.

-35 dB is a reasonable choice. This value matches the acquired S/N for the lowest frequency band when the full observation set is employed and translates to a mask radius $\theta_{\mathrm{mask}} \sim 5\,\theta_{\mathrm{FWHM}}$ for the 93 GHz case. For consistency, we also use $\theta_{\mathrm{mask}} \sim 5\,\theta_{\mathrm{FWHM}}$ for the other frequency bands.

5.5. Dependence on the Frequency Band

Figure 15 shows the reconstructed beam profiles for the four frequency bands compared to their input models. From the figure, we see the beam profiles of the MF and UHF bands following different side-lobe patterns in the target region. The `beamlib` code adapts to these differences by optimizing the interplay between the Bessel function basis model, which fits the beam core (where the side-lobe structure is expected to be more pronounced), and the $1/\theta^3$ fit of the beam wing. The transition from core to wing fit is denoted with a black vertical line. Of all frequency bands, the 280 GHz band has the largest uncertainty on the reconstructed beam profile. The corresponding harmonic transform errors and their uncertainty, as calculated from `beamlib`, are presented in Figure 16. The plot
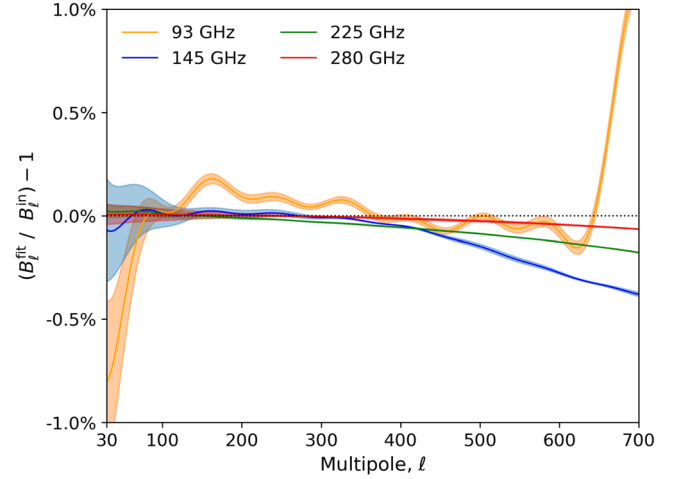
Figure 16. Beam transfer function error with respect to the input beam per frequency band for the beam models of Figure 15.

Table 3
The Reconstructed Beam Transfer Function and Solid Angle Bias for the Different Frequency Bands

Frequency Band [GHz]	δB_ℓ $\ell = 30$–700	δB_ℓ $\ell = 50$–200	$\delta \Omega$
93	$\leqslant 1.2\%$	$\leqslant 0.6\%$	1.8%
145	$\leqslant 0.4\%$	$\leqslant 0.2\%$	0.7%
225	$\leqslant 0.2\%$	$\leqslant 0.06\%$	1.1%
280	$\leqslant 0.1\%$	$\leqslant 0.05\%$	0.7%

Note. The beam transfer function bias is shown for both the full and slightly truncated multipole range.

shows a bias that roughly decreases in amplitude with increasing band center frequency, especially for the multipole range $100 < \ell < 300$, and remains under $\sim 1.3\%$ at all times. Table 3 shows the maximum values of the reconstructed beam transfer function error, $\delta B_\ell = (B_\ell^{\mathrm{fit}}/B_\ell^{\mathrm{in}}) - 1$, for all frequency bands for both the full and a slightly truncated multipole range, which will be further evaluated for calibration against Planck data in Section 5.6.

The multipole region at which B-modes are expected to peak is still well contained within the truncated multipole range. The bias on the solid angle estimation, $\delta \Omega = (\Omega^{\mathrm{fit}}/\Omega^{\mathrm{in}}) - 1$, from `beamlib` is also shown.

These results reflect not only the expected scaling of the S/N as a function of frequency (and associated beam size) but also the success of the basis function choice for the beam model. This argument becomes evident when looking at the ringy pattern of the 93 GHz band transfer function bias and associated uncertainty. Notice that the uncertainty of the reconstructed beam transfer function reduces as the number of available input simulations (and therefore accumulated S/N) increases. An estimate that quantifies this statement is provided in Appendix C.

5.6. Calibration Multipole Range and Error Modes

The technique chosen for the absolute calibration of the beam transfer functions will impact the ℓ-dependence of the bias. Calibrating the SAT beam transfer function against previous CMB experiments, such as Planck, is carried out by matching the spectra of the two telescopes over a limited range

of multipoles. Since B-modes are expected to peak at a multipole number of $\ell \approx 80$, a calibration range around this lower multipole region is naturally motivated.

We test the impact of different calibration choices directly on the bias of the reconstructed beam transfer function compared to the input. We do this by drawing 10^4 realizations, B'_ℓ, of the reconstructed beam transfer function B_ℓ^{fit} and the first 10 error modes $\delta B_\ell^{(i)}$:

$$B'^{j}_\ell = B_\ell^{\text{fit}} + \sum_{i=1}^{10} c_{i,j} \delta B_\ell^{(i)}, \qquad j = 1,...,10^4. \qquad (8)$$

The weights, $c_{i,j}$, are randomly drawn from a normal distribution of zero mean and standard deviation equal to 1.

Given the SAT beams and expected transfer function, the multipole range that facilitates the absolute calibration of the SAT maps using the Planck data will likely lie within $\ell_{\text{min}} = 50$ and $\ell_{\text{max}} = 200$. To investigate the impact of different choices of calibration range, we slice this multipole range and calibrate each one of the beam realizations we produced over the ranges $\ell = 50$–100, 50–200, and 100–200 by minimizing the difference between the output and input beam over each range. Figure 17 shows the 1σ error band of the 10^4 newly calibrated beams, divided by the input beam transfer function, for the three multipole range choices quoted above. The results are shown for all four frequency bands (increasing in frequency from top to bottom) and are truncated to $\ell = 10$–300 for visualization purposes. As one can conclude from the plots, assuming a calibration range of $\ell = 50$–100 results in the minimum beam uncertainty at the low-multipole range of interest, while a calibration range at higher multipoles significantly increases the beam uncertainty at low multipoles.

5.7. Beam Reconstruction Uncertainty Impact on the r-constraint

The SO SATs allow us to constrain the tensor-to-scalar ratio, r, with a statistical error of $\sigma(r) \leqslant 0.003$. Beam modeling errors can bias cosmological analysis, and it is therefore appropriate to briefly consider their impact on the forecasted value of r. For this purpose, we employ the `BBpower` software,[34] which is part of the publicly available SO analysis pipeline (Wolz et al. 2023). `BBpower` is a harmonic-based component separation algorithm that has been adapted to the specifications of the SO telescopes (Simons Observatory Collboration et al. 2019). We use the code to forecast sensitivity to the value of the tensor-to-scalar ratio through Fisher analysis (Fisher 1922).

To quantify the effect of the beam reconstruction bias and uncertainty, we use Equation (8) to create 100 biased beam realizations for each of the four frequency bands considered in our analysis. For each beam realization, we construct a set of beam-convolved CMB and foreground spectra assuming no primordial B-modes ($r = 0$). The sky component and noise power spectra follow "Pipeline A" with the "baseline" noise level and "optimistic" $1/f$ noise description in Wolz et al. (2023). We then forecast the reconstructed r-value for each of the 100 realizations by (incorrectly) using the unbiased input beam to perform beam deconvolution in the Fisher forecast code. The resulting bias on the tensor-to-scalar ratio is $\Delta r = 1.08 \times 10^{-4}$. This number can be compared to the expected 1σ error on r, which is $\sigma(r) \approx 3 \times 10^{-3}$

---

[34] https://github.com/simonsobs/BBPower
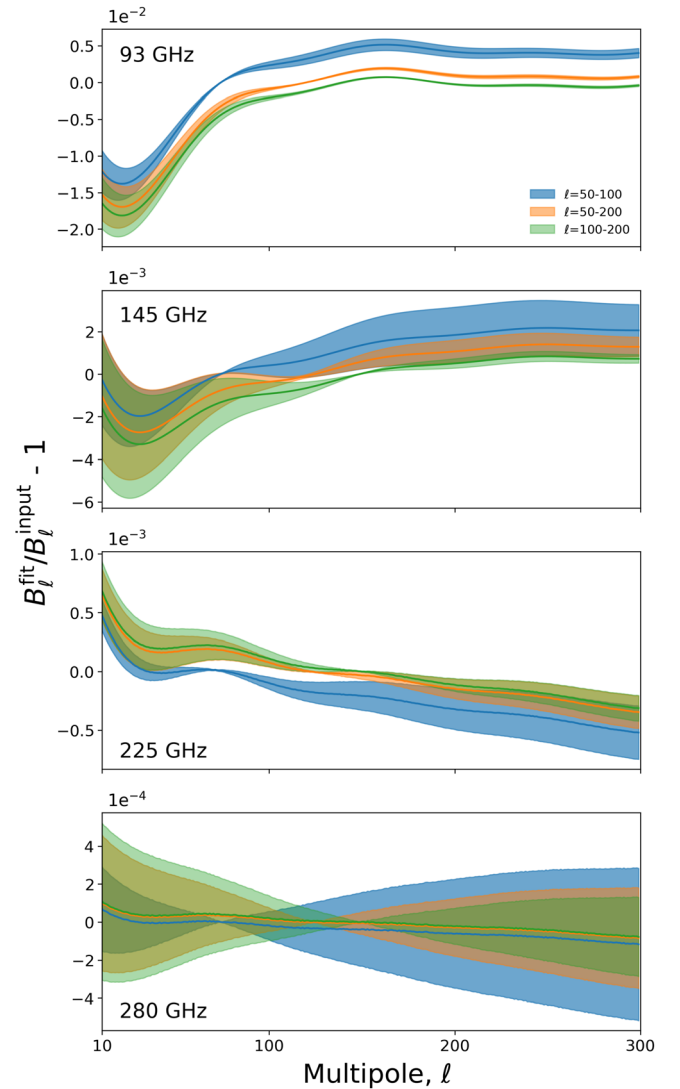
Figure 17. The 1σ band of the beam transfer function bias with respect to the input beam for 10^4 different beam realizations of the best-fit reconstructed beam and first 10 error modes for the four frequency bands centered on 93, 145, 225, and 280 GHz. The beam realization transfer functions are calibrated on three different multipole ranges: $\ell = 50$–100 (blue), 50–200 (orange), and 100–200 (green). All of these ranges are suitable for calibrating the beam transfer function against experiments like Planck.

(Simons Observatory Collboration et al. 2019). These beam errors add insignificantly to the overall variance on r: $\sigma(r)^{\text{extra}} \sim 10^{-6}$. We thus conclude that the beam reconstruction error achieved with the setup presented in this work will be small enough to not significantly bias the SO r-measurement.

6. Conclusions and Discussion

This paper describes a beam reconstruction pipeline for the SO SAT beams in the MF and UHF frequency bands. The low-frequency bands are left for future work. We generate 50 $\sim$1 hr long CES simulations of Jupiter observations (as described in Section 3.3) and feed them to a filter-and-bin mapmaker designed to mitigate the correlated atmospheric noise by removing the strongest modes calculated from a PCA. The maps produced in this way are inputs to a slightly modified version of the ACT beam fitting code, `beamlib`. From this code, we obtain the best-fit beam profiles, transfer functions,

and associated error modes. We present results that quantify the success of our beam fitting method as a function of different input beam models, weather, and frequency bands. These simulations allow us to assess the overall robustness of our analysis pipeline and prepare for the arrival of real data.

Our simulations for the 93 GHz band show that the beam reconstruction is generally robust to optical effects caused by detector location on the focal plane; we are able to reconstruct beam transfer functions with errors not exceeding 1.5% in the $\ell = 30$–700 range and better than 0.6% in the $\ell = 50$–200 range. Testing how beam reconstruction for the 93 GHz band depends on weather parameters shows similar results. This indicates that planet observations are useful even under relatively adverse weather conditions.

The fitted beam profiles and transfer functions vary as a function of frequency. We model all four frequency bands to at least ~ -35 dB and estimate the transfer function bias. The results show the fitting model adapting well to the different side-lobe patterns for the MF and UHF bands and the beam transfer function bias decreasing with increasing frequency. The uncertainty in the beam reconstruction can be reduced by optimizing the range of ℓ used to calibrate the data by comparing it to previous experiments (see Section 5.6). We find the preferred multipole range to be $\ell = 50$–100, as it provides the lowest uncertainty on the beam transfer function over the $\ell = 10$–300 region.

We note that in the beam reconstruction error analysis, marginalization over ad hoc choices, such as the wing scale and the number of subtracted modes, was not included, and that this is different from what was done in Lungu et al. (2022). We expect these sources of error to somewhat increase the beam reconstruction uncertainty, particularly in the low-ℓ regime, but leave this analysis for future work.

Using simulated planet observations with a realistic atmospheric component, we observe beam reconstruction biases that are nonnegligible compared to the uncertainty estimates (see Section 5). However, these multiplicative biases are still relatively small ($<0.6\%$ in the $\ell = 50$–200 range for all cases we tested) and are not expected to significantly impact the cosmological analysis. To verify this, we used a Fisher analysis to propagate the beam reconstruction bias and uncertainty. The result indicates that the reconstruction bias will be small enough to not significantly bias the SO r-constraint.

Acknowledgments

Appendix A
Varying Input Beam

A.1. Passband Variations

Variations in the detector passband will impact the shape of the effective beam and therefore affect the performance of the fitting algorithm. While we leave the detailed analysis of this phenomenon for future work, it is useful to show how the input beam models may change under the assumption of nonuniform passbands. As stated in Section 2, the frequency band beams are produced by combining five monochromatic beams within a 20% bandwidth around the center frequency with a top-hat passband. We compare the profile of the beams that were constructed this way to the ones where, instead of a top-hat function, we employed the simulated passbands from Abitbol et al. (2021) and show the results in Figure 18 for all frequency bands.

Any difference between the uniformly and nonuniformly weighted beam profiles is negligible, at least to the ~ -35 dB level we have chosen for fitting the SAT beams. Consequently, there is no indication from these plots that any change to the beam fitting method would be necessary. The passband assumptions/simulations we make will eventually be replaced with Fourier transform spectroscopy measurements to characterize the instrument's spectral response. These measurements will enable us to produce realistic SAT beam models for future analysis.

A.2. Beam Chromaticity

The instrumental beam can be frequency-scaled in a way that matches the spectral energy distribution (SED) of the different sky components that the telescope observes. Properly accounting for this effect is important for the performance of foreground component separation algorithms. Assuming a known passband, $\tau(\nu)$, of the instrument, the frequency-averaged beam, $B(\theta, \phi)$, can be described as

$$B(\theta, \phi) = \int B(\theta, \phi, \nu)\tau(\nu)S(\nu)d\nu, \qquad (A1)$$

where $B(\theta, \phi, \nu)$ is a monochromatic beam at frequency ν, and $S(\nu)$ captures the assumed frequency scaling. For many astrophysical sources, the latter can be expressed as a power law,

$$S(\nu) = \left(\frac{\nu}{\nu_c}\right)^{\beta}, \qquad (A2)$$
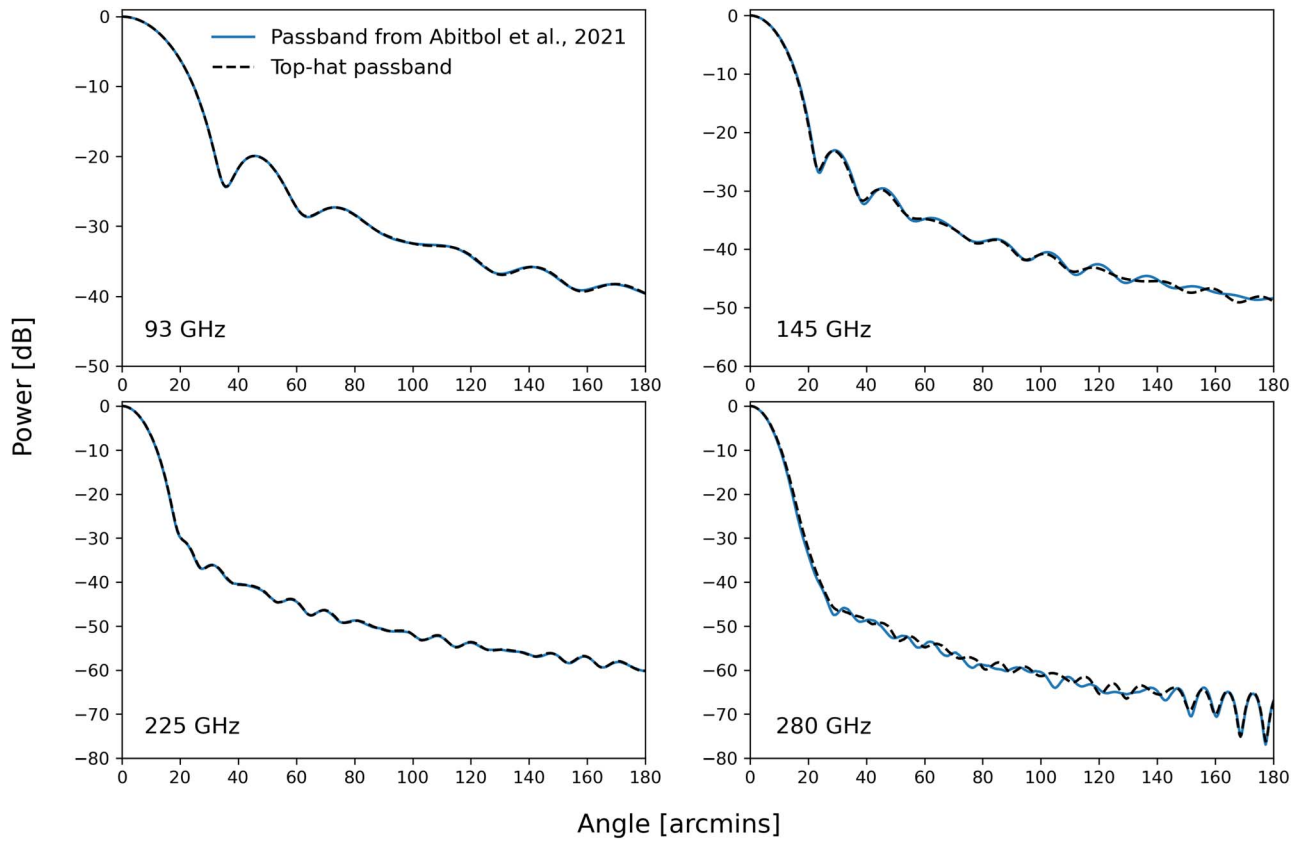
Figure 18. Logarithmic profiles of beam models where a top-hat (black dashed line) and a more realistic (blue line) passband were assumed. The realistic passband was taken from Abitbol et al. (2021), and the results refer to four frequency bands centered on 93, 145, 225, and 280 GHz.

where ν_c is the frequency band center, and β is the spectral index. We consider four cases of frequency scaling matching the SED of the CMB, planets, galactic dust, and synchrotron emission, corresponding to spectral index values of $\beta_{\mathrm{CMB}} = 1$, $\beta_{\mathrm{planet}} = 2$, $\beta_{\mathrm{dust}} = 1.56$, and $\beta_{\mathrm{sync}} = -3$ (Planck Collaboration et al. 2020). The beam profiles for these four cases, along with the case where no frequency scaling was implemented, are shown in Figure 19 for all four frequency bands.

It is interesting to see how the frequency scaling impacts the beam transfer function. Figure 20 shows the ratio of the beam transfer function of the four chromatic beams described above and the one of the nominal case where no frequency scaling was implemented. The chromaticity effect is smooth across all frequency bands and decreases in amplitude with increasing frequency. In the case of the 93 GHz band, not taking account of the beam frequency scaling can result in a transfer function bias as large as ~10% at $\ell = 700$.
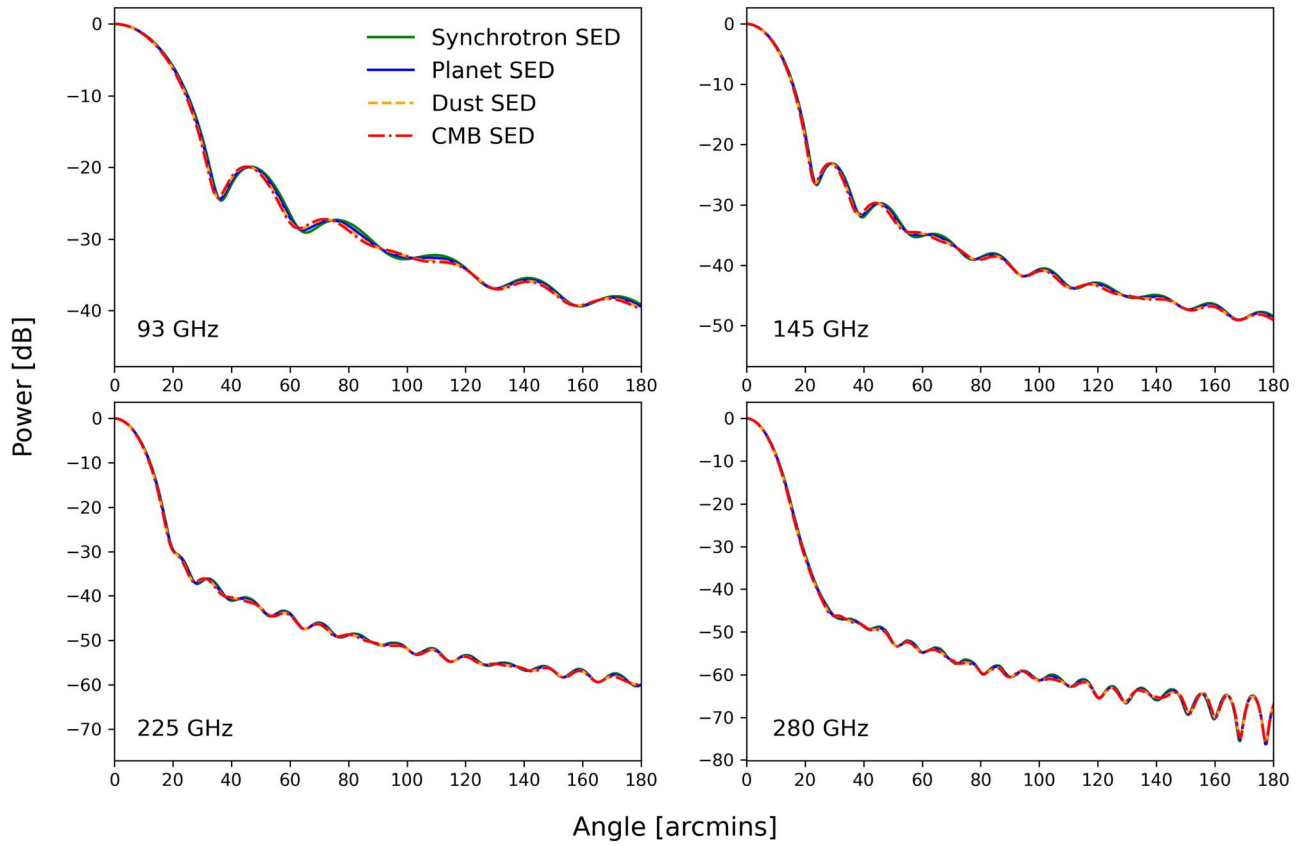
Figure 19. Beam profiles for four models constructed using Equations (A1) and (A2), where the frequency scaling matches the SED of the CMB (red curve), planets (blue curve), galactic dust (orange curve), and synchrotron emission (green curve). The plots refer to four frequency bands centered on 93, 145, 225, and 280 GHz and include the case where no frequency scaling was implemented (black dashed line) for reference.
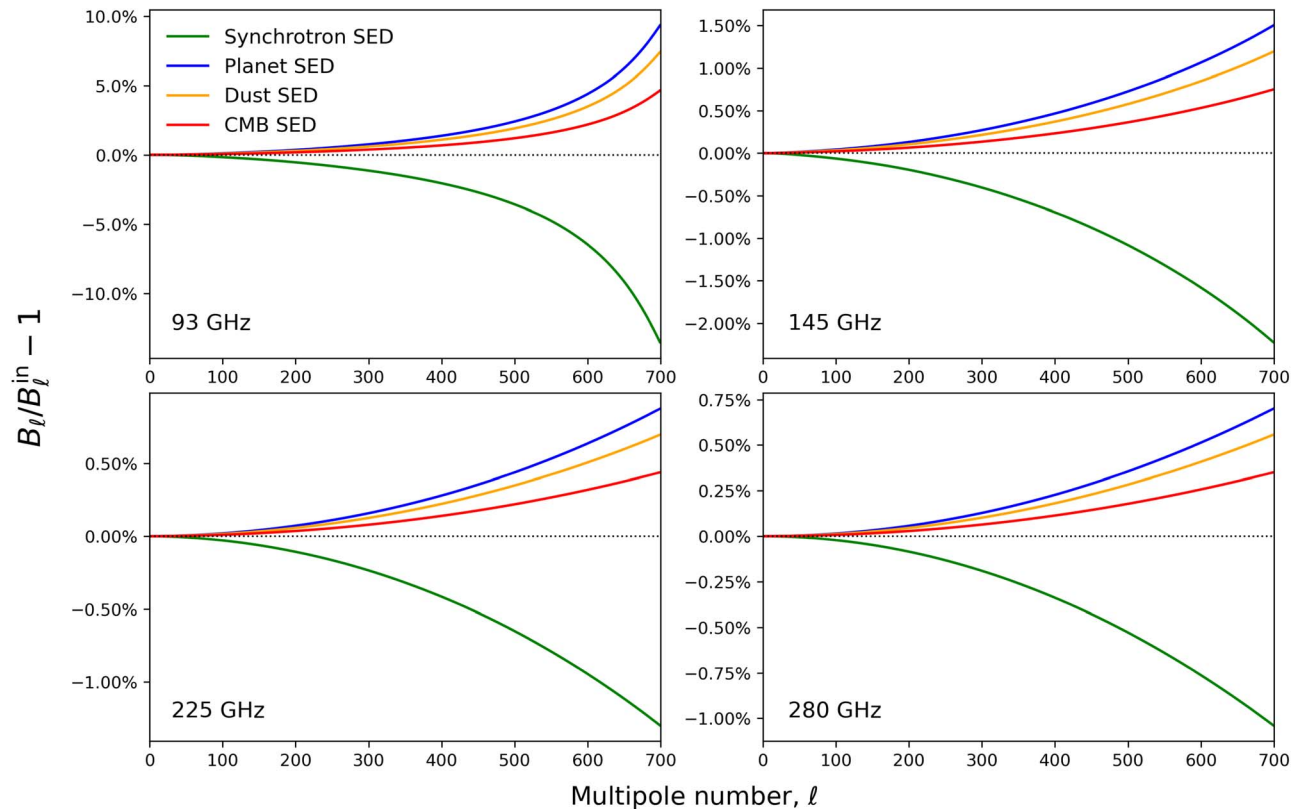
Figure 20. The ratio of the beam transfer functions for the four chromatic beams whose profiles were shown in Figure 19 and the transfer function of the nominal case where no frequency scaling was implemented. The results are shown for four frequency bands centered on 93, 145, 225, and 280 GHz.

Appendix B
Signal Strength Estimation for Different Sources

The total power (in watts) received by a radio telescope due to an astrophysical source can be expressed as

$$P_{\text{received}} = \iint d\Omega d\nu \tau'(\nu) A_{\text{eff}}(\nu) B(\theta, \phi, \nu) S(\theta - \theta_0, \phi - \phi_0, \nu, T), \quad \text{(B1)}$$

where $A_{\text{eff}}(\nu)$ is the telescope effective area, $B(\theta, \phi, \nu)$ is the frequency-dependent beam response of the telescope, $S(\theta - \theta_0, \phi - \phi_0, \nu, T)$ captures the SED of the source parameterized using the Planck blackbody equation and thermodynamic temperature T, and $\tau'(\nu)$ captures the spectral response function of the telescope, including effects from the finite transmissivity of the Earth's atmosphere. The planets' thermodynamic temperatures were taken from Planck Collaboration LII et al. (2017), and we have used $S(\nu, T)$ to represent the Planck blackbody formula for spectral radiance instead of $B(\nu, T)$ to prevent confusion with the beam response.

If the source is small relative to the size of the telescope's beam response subtended on the sky, we can collapse the solid angle convolution and write

$$P_{\text{received}} \approx \frac{\Omega_{\text{source}}}{\Omega_{\text{beam}}} \int S(\nu, T)\tau(\nu)d\nu \approx A_{\text{eff}} \Omega_{\text{source}} \int S(\nu, T)\tau'(\nu)d\nu, \quad \text{(B2)}$$

where we have assumed that the spectral response function, $\tau(\nu)$, can be written as

$$\tau(\nu) \equiv \tau'(\nu) \cdot n\lambda^2 = \tau'(\nu) \cdot A_{\text{eff}}(\nu)\Omega_{\text{beam}}(\nu), \quad \text{(B3)}$$

with n corresponding to the number of radiation modes (Hudson 1974; Hodara & Slemon 1984). In Equation (B2), we have made the approximation that the frequency dependence of the beam solid angle, $\Omega_{\text{beam}}(\nu)$, or, correspondingly, the telescope effective area, $A_{\text{eff}}(\nu)$, can be ignored. The accuracy of this approximation depends, of course, on the width of the frequency range over which we must integrate.

For a restricted range of frequencies centered on $\nu = \nu_0$, the above equation can be further simplified to

$$P_{\text{received}} \approx A_{\text{eff}} \Omega_{\text{source}} \cdot S(\nu_0, T)\tau'(\nu_0)\Delta\nu, \quad \text{(B4)}$$

where $\Delta\nu$ corresponds to the frequency bandwidth over which the signal from the source is integrated, and ν_0 is the band center frequency. In the above equation, $\tau'(\nu_0)$ describes the telescope's spectral response function that ignores the $n\lambda^2$ scaling (see Equation (B3)). Note, however, that losses in signal strength due to instrumental effects such as nonideal optical efficiency should be included in $\tau'(\nu_0)$. From Equation (B4), it is clear that we can calculate the signal strength from a compact source, assuming that we know all of the parameters on the right-hand side of the equation. We use this equation to calculate the signal amplitude from Jupiter, Mars, and Saturn.

The power incident on a telescope from an artificial source has been estimated using the Friis equation (Friis 1946; Johnson 1993),

$$P_{\text{incident}} = \left(\frac{P_{\text{emit}} \cdot A_{\text{tel}}}{4\pi d^2}\right)10^{g/10}, \quad \text{(B5)}$$

where P_{emit} is the in-band power emitted by the source, $g = 10\log_{10}(4\pi/\Omega_{\text{ant}})$ is the forward gain of the antenna, and d is the distance from the source. Note that this expression can be extended to explicitly include integration over the spectral bandpass, but for simplification, it is common to assume that all power emitted by the source is in the spectral band of the receiver.

As in the case of power from a compact astrophysical source, the power arriving on the detector element must account for attenuation due to the optical elements between the outside of the telescope and the focal plane itself. We therefore write

$$P_{\text{received}} = \eta P_{\text{incident}}, \quad \text{(B6)}$$

where η describes the optical efficiency of the telescope. The atmosphere is considered transparent for the artificial source ($\mathcal{T} \simeq 1$) given the simulated frequencies and the short distance of the artificial source, $d \simeq 500$ m.

Finally, the S/N from a compact astrophysical source or a transmitter mounted on a drone is estimated as

$$\text{S/N} = \frac{P_{\text{received}}}{\text{NEP} \cdot \sqrt{2/t}}, \quad \text{(B7)}$$

where NEP represents the assumed noise-equivalent power, and t is the integration time per pixel, which we have set to 1 s for generating numbers for Figure 3.

Appendix C
Beam Transfer Function Uncertainty as a Function of the Number of Observations

The scaling of the beam fitting performance with the number of input simulations needs to be quantified. From Figure 14, we conclude an increase in the S/N by about half an order of magnitude when increasing the number of available observations by a factor of 10. Ideally, we would like the benefit of increasing the input number of observations to be strongly pronounced on the beamlib results as well.

We test this by running the beam fitting code for the same parameters used for the results shown in Figures 15 and 16 but alter the number of input observations we provide to the code each time. This number ranges again from 5 to 50 observations. We construct distributions of 10^3 samples of the 10 strongest error modes of the recovered beam transfer function as calculated from beamlib, and we estimate the standard deviation of these distributions per observation number and frequency band.

Figure 21 shows the fractional standard deviation of the beam transfer function uncertainty distribution with respect to the best-fit estimation as a function of the observation number for four frequency bands centered on 93, 145, 225, and 280 GHz. The uncertainty decreases by a factor of $\gtrsim 2$ when increasing the number of input observations we feed to the code by a factor of 10. This result is consistent for all four frequency bands.
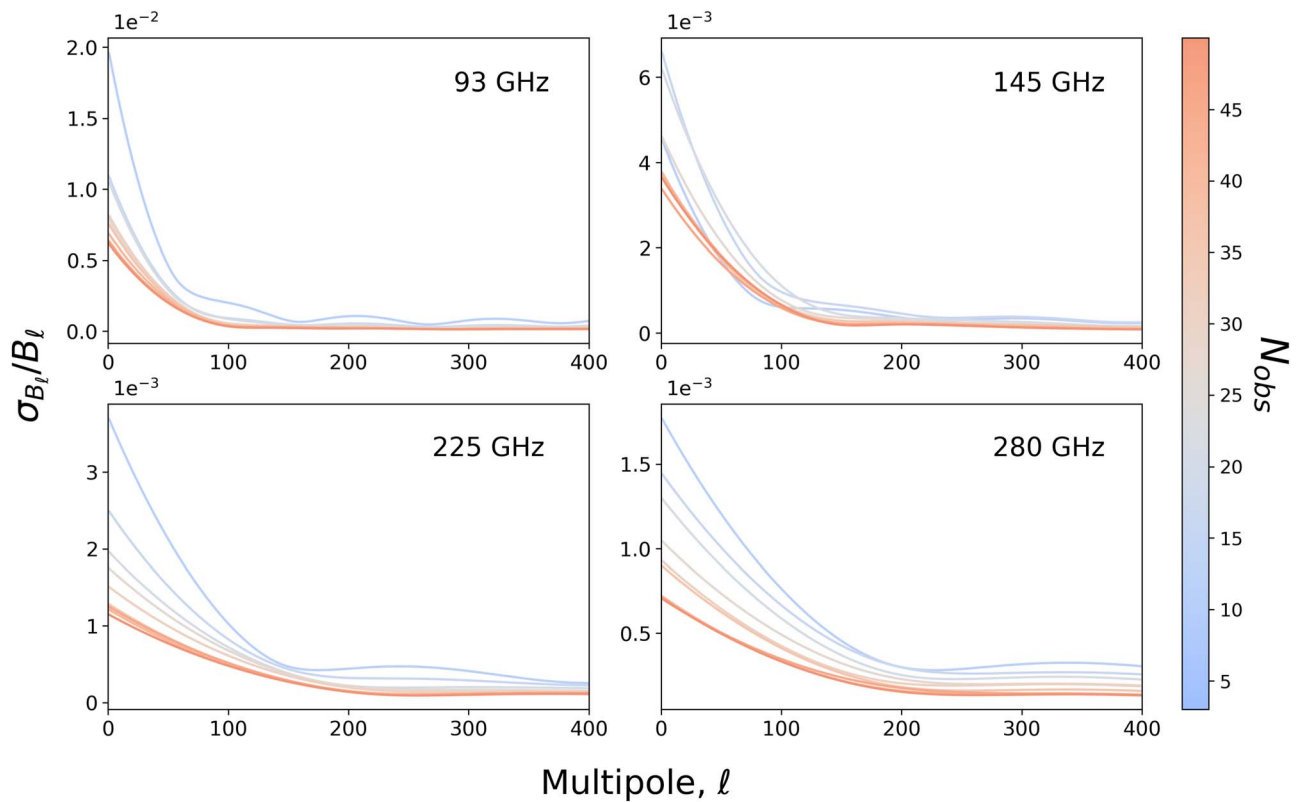
Figure 21. Standard deviation of the beam transfer function uncertainty distribution, as calculated from the 10 strongest error modes, divided by the best-fit transfer function value. The standard deviation is plotted vs. the number of input observations provided to `beamlib` for four frequency bands centered on 93, 145, 225, and 280 GHz.

ORCID iDs

Nadia Dachlythra https://orcid.org/0009-0006-7382-1434
Adriaan J. Duivenvoorden https://orcid.org/0000-0003-2856-2382
Jon E. Gudmundsson https://orcid.org/0000-0003-1760-0355
Matthew Hasselfield https://orcid.org/0000-0002-2408-9201
Gabriele Coppi https://orcid.org/0000-0002-6362-6524
Alexandre E. Adler https://orcid.org/0000-0002-5736-5524
David Alonso https://orcid.org/0000-0002-4598-9719
Susanna Azzoni https://orcid.org/0000-0002-8132-4896
Grace E. Chesmore https://orcid.org/0000-0001-6702-0450
Giulio Fabbian https://orcid.org/0000-0002-3255-4695
Ken Ganga https://orcid.org/0000-0001-8159-8208
Remington G. Gerras https://orcid.org/0009-0009-0876-9168
Andrew H. Jaffe https://orcid.org/0000-0003-2086-1759
Bradley R. Johnson https://orcid.org/0000-0002-6898-8938
Brian Keating https://orcid.org/0000-0003-3118-5514
Reijo Keskitalo https://orcid.org/0000-0001-5748-5182
Theodore S. Kisner https://orcid.org/0000-0003-3510-7134
Nicoletta Krachmalnicoff https://orcid.org/0000-0002-5501-8449
Marius Lungu https://orcid.org/0000-0002-3800-5558
Frederick Matsuda https://orcid.org/0000-0003-0041-6447
Sigurd Naess https://orcid.org/0000-0002-4478-7111
Lyman Page https://orcid.org/0000-0002-9828-3525
Giuseppe Puglisi https://orcid.org/0000-0002-0689-4290
Sara M. Simon https://orcid.org/0000-0001-9221-7802
Edward J. Wollack https://orcid.org/0000-0002-7567-4451

Kevin Wolz https://orcid.org/0000-0003-3155-6151
Zhilei Xu https://orcid.org/0000-0001-5112-2567

References

Abitbol, M. H., Alonso, D., Simon, S. M., et al. 2021, JCAP, 2021, 032
Adler, A., & Gudmundsson, J. E. 2020, Proc. SPIE, 11453, 114534O
Aikin, R. W., Ade, P. A., Benton, S., et al. 2010, Proc. SPIE, 7741, 77410V
Ali, A. M., Adachi, S., Arnold, K., et al. 2020, JLTP, 200, 461
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
BICEP2/Keck Array XI, Ade., P. A. R., Ahmed, Z., et al. 2019, ApJ, 884, 114
Choi, S. K., Hasselfield, M., Ho, S.-P. P., et al. 2020, JCAP, 2020, 045
Coppi, G., Conenna, G., Savorgnano, S., et al. 2022, Proc. SPIE, 12190, 1219015
Duivenvoorden, A. J., Adler, A. E., Billi, M., Dachlythra, N., & Gudmundsson, J. E. 2021, MNRAS, 502, 4526
Dünner, R., Fluxá, J., Best, S., & Carrero, F. 2020, Proc. SPIE, 11453, 114532P
Dünner, R., Fluxá, J., Best, S., Carrero, F., & Boettger, D. 2021, in 15th European Conference on Antennas and Propagation (EuCAP), 15 (Piscataway, NJ: IEEE), 1
Dünner, R., Hasselfield, M., Marriage, T. A., et al. 2012, ApJ, 762, 10
Errard, J., Ade, P. A. R., Akiba, Y., et al. 2015, ApJ, 809, 63
Fisher, R. A. 1922, RSPTA, 222, 309
Fraisse, A., Ade, P., Amiri, M., et al. 2013, JCAP, 2013, 047
Friis, H. 1946, PIRE, 34, 254
Galitzki, N. 2018, The Simons Observatory: Project Overview, arXiv:1810.02465
Global Modeling and Assimilation Office (GMAO) 2015, Global Modeling and Assimilation Office (GMAO) 2015, MERRA-2 tavg1_2d_slv_Nx: 2d,1-Hourly,Time-Averaged,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4, https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/citing_MERRA-2/
Hasselfield, M., Moodley, K., Bond, J. R., et al. 2013, ApJS, 209, 17
Hill, C. A., Kusaka, A., Barton, P., et al. 2018, JLTP, 193, 851
Hodara, H., & Slemon, C. 1984, Appl. Sci., 41, 203
Hudson, M. C. 1974, ApOpt, 13, 1029

Johnson, B. R., Vourch, C. J., Drysdale, T. D., et al. 2015, JAI, 4, 1550007

Johnson, R. C. 1993, Antenna Engineering Handbook (3rd ed.; New York: McGraw-Hill)

Keating, B. G., Shimon, M., & Yadav, A. P. S. 2013, ApJL, 762, L23

Kolmogorov, A. 1941, DSSRA, 30, 301

Komatsu, E. 2022, NatRP, 4, 452

Krachmalnicoff, N., Baccigalupi, C., Aumont, J., Bersanelli, M., & Mennella, A. 2016, A&A, 588, A65

Lungu, M., Storer, E. R., Hasselfield, M., et al. 2022, JCAP, 2022, 044

Masi, S., Ade, P. A. R., Bock, J. J., et al. 2006, A&A, 458, 687

Matsuda, F. T. 2020, Proc. SPIE, 11445, 114456Z

Morris, T. W., Bustos, R., Calabrese, E., et al. 2022, PhRvD, 105, 042004

Næss, S. K. 2019, JCAP, 2019, 060

Nati, F., Devlin, M. J., Gerbino, M., et al. 2017, JAI, 6, 1740008

Planck Collaboration 2020, A&A, 641, A4

Planck Collaboration I 2020, A&A, 641, A1

Planck Collaboration VII 2016, A&A, 594, A7

Planck Collaboration LII 2017, A&A, 607, A122

Puglisi, G., Keskitalo, R., Kisner, T., & Borrill, J. D. 2021, RNAAS, 5, 137

Salatino, M., Lashner, J., Gerbino, M., et al. 2018, Proc. SPIE, 10708, 48

Shimon, M., Keating, B., Ponthieu, N., & Hivon, E. 2008, PhRvD, 77, 083003

Simons Observatory Collaboration 2019, JCAP, 2019, 056

Takakura, S., Aguilar-Faúndez, M. A. O., Akiba, Y., et al. 2019, ApJ, 870, 102

Weiland, J. L., Odegard, N., Hill, R. S., et al. 2011, ApJS, 192, 19

Wolz, K., Azzoni, S., Hervias-Caimapo, C., et al. 2023, arXiv:2302.04276

Xu, Z., Bhandarkar, T., Coppi, G., et al. 2020, Proc. SPIE, 11453, 1145315