

# The Atacama Cosmology Telescope: semi-analytic covariance matrices for the DR6 CMB power spectra

Zachary Atkins <sup>a,\*</sup> Zack Li <sup>b,c,d</sup> David Alonso <sup>e</sup> J. Richard Bond <sup>d</sup>  
Erminia Calabrese <sup>f</sup> Adriaan J. Duivenvoorden <sup>g,a,h</sup> Jo Dunkley <sup>a,i</sup>  
Serena Giardiello <sup>f</sup> Carlos Hervías-Caimapo <sup>j</sup> J. Colin Hill <sup>k</sup> Hidde T. Jense <sup>f</sup>  
Joshua Kim <sup>l</sup> Thibaut Louis <sup>m</sup> Kavilan Moodley <sup>n,o</sup> Thomas W. Morris <sup>p,q</sup>  
Sigurd Naess <sup>r</sup> Michael D. Niemack <sup>s,t</sup> Lyman Page <sup>a</sup> Adrien La Posta <sup>e</sup>  
Cristóbal Sifón <sup>u</sup> and Edward J. Wollack <sup>v</sup>

<sup>a</sup>Joseph Henry Laboratories of Physics, Jadwin Hall, Princeton University,  
Princeton, NJ 08544, U.S.A.

<sup>b</sup>Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.

<sup>c</sup>Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, U.S.A.

<sup>d</sup>Canadian Institute for Theoretical Astrophysics, University of Toronto, McLennan Labs,  
Toronto, ON, M5S 3H8, Canada

<sup>e</sup>Department of Physics, University of Oxford, Denys Wilkinson Building,  
Keble Road, Oxford OX1 3RH, U.K.

<sup>f</sup>School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, Wales CF24 3AA, U.K.

<sup>g</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany

<sup>h</sup>Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, U.S.A.

<sup>i</sup>Department of Astrophysical Sciences, Peyton Hall, Princeton University,  
Princeton, NJ 08544, U.S.A.

<sup>j</sup>Instituto de Astrofísica and Centro de Astro-Ingeniería, Facultad de Física,  
Pontificia Universidad Católica de Chile,  
Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile

<sup>k</sup>Department of Physics, Columbia University, New York, NY 10027, U.S.A.

<sup>l</sup>Department of Physics and Astronomy, University of Pennsylvania,  
209 South 33rd Street, Philadelphia, PA 19104, U.S.A.

<sup>m</sup>Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

<sup>n</sup>Astrophysics Research Centre, University of KwaZulu-Natal,  
Westville Campus, Durban 4041, South Africa

<sup>o</sup>School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal,  
Westville Campus, Durban 4041, South Africa

<sup>p</sup>Department of Physics, Yale University, New Haven, CT 06511, U.S.A.

<sup>q</sup>National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, U.S.A.

\*Corresponding author.

<sup>r</sup>*Institute of Theoretical Astrophysics, University of Oslo, Norway*

<sup>s</sup>*Department of Physics, Cornell University, Ithaca, NY 14853, U.S.A.*

<sup>t</sup>*Department of Astronomy, Cornell University, Ithaca, NY 14853, U.S.A.*

<sup>u</sup>*Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile*

<sup>v</sup>*NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, U.S.A.*

*E-mail:* [zatkins@princeton.edu](mailto:zatkins@princeton.edu)

ABSTRACT: The Atacama Cosmology Telescope Data Release 6 (ACT DR6) power spectrum is expected to provide state-of-the-art cosmological constraints, with an associated need for precise error modeling. In this paper we design, and evaluate the performance of, an analytic covariance matrix prescription for the DR6 power spectrum that sufficiently accounts for the complicated ACT map properties. We use recent advances in the literature to handle sharp features in the signal and noise power spectra, and account for the effect of map-level anisotropies on the covariance matrix. In including inhomogeneous survey depth information, the resulting covariance matrix prescription is structurally similar to that used in the *Planck* Cosmic Microwave Background (CMB) analysis. We quantify the performance of our prescription using comparisons to Monte Carlo simulations, finding better than 3% agreement. This represents an improvement from a simpler, pre-existing prescription, which differs from simulations by  $\sim 16\%$ . We develop a new method to correct the analytic covariance matrix using simulations, after which both prescriptions achieve better than 1% agreement. This correction method outperforms a commonly used alternative, where the analytic correlation matrix is assumed to be accurate when correcting the covariance. Beyond its use for ACT, this framework should be applicable for future high resolution CMB experiments including the Simons Observatory (SO).

KEYWORDS: CMBR experiments, cosmological parameters from CMBR, Statistical sampling techniques

ARXIV EPRINT: [2412.07068](https://arxiv.org/abs/2412.07068)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data and deliverables</b>	<b>2</b>
<b>3</b>	<b>Analytical framework</b>	<b>4</b>
3.1	ACT data model	4
3.2	Covariance matrices in MASTER	8
3.3	Approximate data model	9
<b>4</b>	<b>Covariance matrix pipeline</b>	<b>10</b>
4.1	Analytic covariance matrix	10
4.2	Simulations	16
4.3	Simulation-based correction	17
<b>5</b>	<b>Results</b>	<b>19</b>
5.1	Covariance matrix pipeline validation	19
5.2	Monte Carlo matrix validation	23
5.3	Discussion of semi-analytic matrix performance	24
<b>6</b>	<b>Conclusion</b>	<b>26</b>
<b>A</b>	<b>More details on data model</b>	<b>28</b>
<b>B</b>	<b>Analytical pseudospectrum covariance matrices</b>	<b>30</b>
B.1	Covariance matrices	30
B.2	More details on Fourier filter	31
B.3	Parameters for fiducial signal spectra	35
<b>C</b>	<b>Updates for <code>mmms</code> noise simulations</b>	<b>35</b>
C.1	Excess large-scale polarization noise power in simulations	37
<b>D</b>	<b>Effect of smoothing diagonals of <math>\beta \neq \beta'</math> blocks of <math>\Sigma_R</math></b>	<b>37</b>
<b>E</b>	<b>Effect of point-source holes on the covariance matrix</b>	<b>38</b>

---

**1 Introduction**

The Atacama Cosmology Telescope’s (ACT) [28, 37, 69] sixth data release (DR6) promises to provide competitive constraints on cosmological models. The measured primary CMB anisotropy power spectrum will be used to test the  $\Lambda$ CDM model and to look for evidence for beyond- $\Lambda$ CDM physics, including for example an early dark energy (EDE) component [38] or self-interacting neutrinos [45]. The characterization of the power spectrum covariance

matrix is critical to these efforts. The pseudo- $C_\ell$  Monte Carlo Apodized Spherical Transform Estimator (MASTER) [39] formalism provides a nearly optimal, unbiased method for power spectrum reconstruction [24]. A key feature is its ability to correct for a common real-world systematic effect — the incomplete sky coverage of surveys — via the computation of “mode-coupling” matrices. Moreover, it is computationally tractable: the computation of these matrices scales with survey resolution as  $\mathcal{O}(\ell_{\text{max}}^3)$  [24, 53], the same scaling as the spherical harmonic transform (SHT) [66]. For these reasons, MASTER has proven versatile, providing the basis for an array of results from CMB analyses [e.g., 9, 16, 23, 26, 47, 63], as well as clustering of galaxies and quasars, and weak lensing [e.g., 2, 21, 22, 27, 60]. The ACT DR6 power spectrum pipeline uses the MASTER formalism.

It has long been recognized that a direct calculation of the MASTER covariance is intractable [see e.g., 11, 14, 24]. In particular, incomplete survey coverage couples to spatial correlations in the underlying field of interest; a fast approximation (that is, also  $\mathcal{O}(\ell_{\text{max}}^3)$  complexity) is only possible assuming the coverage is sufficiently uniform [see e.g., 18, 30]. Analyses utilizing the MASTER formalism have thus investigated the validity of this assumption for their specific application [e.g., 29, 62]. The community continues to develop approaches to handle non-uniform survey geometry more robustly [e.g., 12, 59]. The ACT DR6 data present a new challenge: the atmosphere induces stronger noise correlations than are present in space-based microwave observatories such as *Planck* or galaxy surveys with close-to-Poisson noise. The data are not adequately described by the simple model required by the MASTER covariance formalism.

In this paper, we investigate the performance of a MASTER-compatible covariance matrix for the ACT DR6 power spectrum. The covariance structure resembles that of ref. [63] in including inhomogeneous survey depth, and uses the analytical framework presented in ref. [59]. We also develop a new method for handling anisotropies introduced in the Fourier-space filtering of the ACT data. We test this covariance matrix alongside a commonly-used analytic version that assumes homogeneous survey depth, and find that the former performs better as measured by agreement with a Monte Carlo covariance matrix. Lastly, we develop a new procedure using the Monte Carlo covariance to apply a smooth correction to either matrix version and achieve sub-percent agreement with simulations.

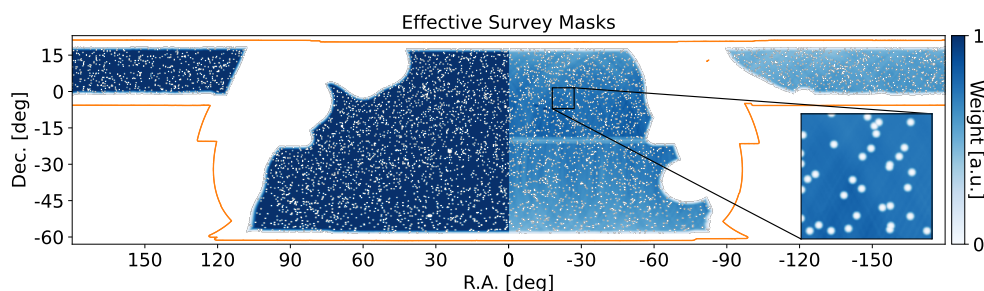
The outline of this paper is as follows: we introduce the ACT DR6 data, and the power spectrum covariance matrix products, in section 2. In section 3, we discuss the MASTER framework, focusing on the particular complications of the ACT data. An overview of our covariance matrix pipeline is provided in section 4, and we assess the results of that pipeline in section 5. We conclude in section 6.

## 2 Data and deliverables

Between 2007 and 2022, ACT observed the microwave sky from Cerro Toco in the Atacama Desert, Chile. ACT DR6 comprises the data collected by the Advanced ACT receiver from 2017 onwards [15, 20, 37, 40, 50]; the power spectrum analysis uses the exclusively-nighttime data (between 23:00 and 11:00 UTC) from three dichroic, polarization-sensitive detector arrays abbreviated “PA4,” “PA5,” and “PA6.” Their beam and approximate observed frequencies — labeled by a frequency “band” — are given in table 1. For each frequency band on each

Array	Band	Frequencies (GHz)	Beam (arcmin)
PA4	f150	124–172	1.4
	f220	182–277	1.0
PA5	f090	77–112	2.0
	f150	124–172	1.4
PA6	f090	77–112	2.0
	f150	124–172	1.4

**Table 1.** Frequency coverage (the 0.5% and 99.5% locations of their cumulative bandpower) and resolution (beam full-width half-maximum in arcminutes) of the Advanced ACT detector arrays used in DR6. Reproduced from [7]. The Author(s). CC BY 4.0.



**Figure 1.** *Blue, Left:* the power spectrum pipeline analysis mask for PA6 f150. *Blue, Right:* the same mask after including effective noise weights (arbitrarily normalized), described in section 4.1. The inset provides a zoomed-in view of the point-source holes. The outer mask borders (the point-source holes) have a  $2^\circ$  ( $0.3^\circ$ ) cosine apodization. *Orange:* the outline of the ACT survey footprint. Data within the orange outline, but not highlighted in blue, are excluded from the analysis.

array, ACT maps four disjoint “splits” of the raw data — formed by separately allocating each day of observations to a given split — such that the noise in each split is independent. The maps are produced for each of the Stokes I, Q, and U polarization components in CMB blackbody temperature units. The power spectrum analysis, including this paper, uses an updated map version, “dr6.02,” which features lower noise and reduced systematics compared to the dr6.01 map version used in refs. [7, 19, 65].

Accompanying the data maps are “inverse-variance maps” [see e.g. 1, 56], which give the white-noise level of the data. In section 4, we use these to derive maps of the white-noise standard deviation per pixel by taking their inverse square-root. The DR6 power spectrum pipeline also includes analysis masks which are applied to the maps, with an example shown in figure 1. The masks have apodized boundaries to limit mode-coupling, and avoid both the edge of the ACT survey footprint where data are noisiest as well as the plane of the Galaxy where foregrounds are brightest. The masks also contain  $\sim 10,000$  apodized point-source holes with a diameter of five arcmin. These correspond to a catalog of sources with measured flux greater than 15 mJy in the 150 GHz channel and detected with a significance greater than  $5\sigma$ .

As will be described in forthcoming papers, the ACT DR6 power spectrum analysis includes substantial data validation that took place with blinded results, resulting in the

removal of PA4 f150 (temperature and polarization) and PA4 f220 polarization from the analysis. Scale cuts are applied to the remaining data that range from a minimum angular multipole of 475–975, depending on array and polarization. The maximum multipole is  $\ell_{\max} = 8,500$  for all arrays. The spectra are binned following the same schedule as presented in ref. [16], appendix F. Each polarized cross spectrum contains up to 57 bins, and the entire data vector including all  $T$ ,  $E$ , and  $B$  spectra, has over 3,500 elements. The baseline DR6 likelihood only uses  $TT$ ,  $TE$ , and  $EE$  spectra; after the cuts we use in this paper, the data vector has 1,763 elements.<sup>1</sup>

A “first version” covariance matrix for the ACT power spectrum was constructed using pre-existing tools available in `pspy`<sup>2</sup> and `PSPipe`.<sup>3</sup> We label this matrix the “homogeneous” matrix for reasons discussed in section 4. In this paper we develop a new covariance matrix; we label it the “inhomogeneous” matrix. Both versions start with an “analytic” covariance matrix, which is then corrected using simulations. For clarity, we refer to the post-correction matrix as the “corrected” or “semi-analytic” matrix. This semi-analytic matrix is the product used in the likelihood to constrain cosmological models. The homogeneous matrix is the default version used in the baseline DR6 likelihood, as it was developed earlier, but as we show in this paper, we find that our new inhomogeneous version requires a smaller simulation-based correction — in other words, that its analytic prescription better describes the data covariance.

### 3 Analytical framework

In this section we describe a model for the ACT DR6 data, as well as approximations that simplify the construction of the power spectrum covariance matrix.

#### 3.1 ACT data model

The ACT data contain several properties that, in principle, are challenging to incorporate into the MASTER framework. We start by describing a generative model for the ACT data units: the 20 polarized split maps enumerated in section 2. Most ACT map metadata are bundled into an “array” label, but for the formalism it is useful to keep the map polarization and split index separate. Thus, we label a given map,  $\mathbf{m}$ , as  $\mathbf{m}_{I_i}^X$ , where  $I$  denotes the array,  $i$  the split index, and  $X$  the polarization component. We consider the map to be a sum of signal and noise components:

$$\mathbf{m}_{I_i}^X = \mathbf{s}_I^X + \mathbf{n}_{I_i}^X \quad (3.1)$$

where the signal does not depend on the map split.<sup>4</sup>

<sup>1</sup>Minor revisions were made to the dr6.02 maps, analysis masks, scale cuts, fiducial signal spectra (section 4.1.3), and simulations (section 4.2) after the completion of this paper, but should not affect our conclusions.

<sup>2</sup><https://github.com/simonsobs/pspy>.

<sup>3</sup><https://github.com/simonsobs/PSPipe>.

<sup>4</sup>This is true insofar as the beam and passband are the same for each split of an array. Beam variations over split in dr6.02 maps are  $\mathcal{O}(0.1\%)$  at signal-dominated scales and so we neglect them. There is no evidence for split-dependent passband variations.

We assume the sky signal to be a realization from a homogeneous, isotropic Gaussian distribution that is subsequently modified by the instrument.<sup>5</sup> Its power spectrum includes contributions from the primary CMB, as well as secondary effects including CMB lensing, the thermal and kinetic Sunyaev Zeldovich effects (tSZ and kSZ), the cosmic infrared background (CIB), unresolved (“point source”) active galactic nuclei (AGN), and dust. The sky signal is passed through the telescope, receiver optics, and filtering, and is sampled by the detector arrays at the focal plane. This process integrates the frequency dependence of each component over each array’s spectral passband and spatially convolves the continuous sky with the instrumental beam.<sup>6</sup> We assume transient instrumental and astrophysical sources can be modeled by their time-averaged quantities over the course of ACT observations.

Collectively, we have the following model of the signal:

$$\mathbf{s} = \mathbf{YBS}^{\frac{1}{2}}\boldsymbol{\eta}_s^h, \quad (3.2)$$

where we have suppressed the array ( $I$ ) and polarization ( $X$ ) labels for simplicity. We read from right to left, starting with  $\boldsymbol{\eta}_s^h$ , which is a white-noise vector in harmonic space. Thereafter, each symbol is a linear operator that acts on the vector:  $\mathbf{S}$  is the power spectrum of the sky signal with elements that covary over arrays (according to their passbands) and polarizations,<sup>7</sup>  $\mathbf{B}$  is the isotropic beam transfer function<sup>8</sup> for each array and polarization, and  $\mathbf{Y}$  is a spherical harmonic transform (SHT) “synthesis” operation (i.e., the matrix that transforms vectors from harmonic space to map space). Note that both  $\mathbf{S}^{\frac{1}{2}}$  and  $\mathbf{B}$  are assumed to be diagonal over spherical harmonic modes, such that this model of the map signal is, like the underlying sky, homogeneous, isotropic, and Gaussian.

Like the signal, the map-level noise is Gaussian distributed, but is inhomogeneous and anisotropic: the noise power varies over the sky, and its correlation structure is “stripy.” Moreover, the noise anisotropy is itself inhomogeneous: the stripy correlation pattern varies over the sky (see section 3 of ref. [7] for a complete description, and figure 2 for illustration). These properties are the product of projecting correlated atmospheric noise along the ACT scanning direction.

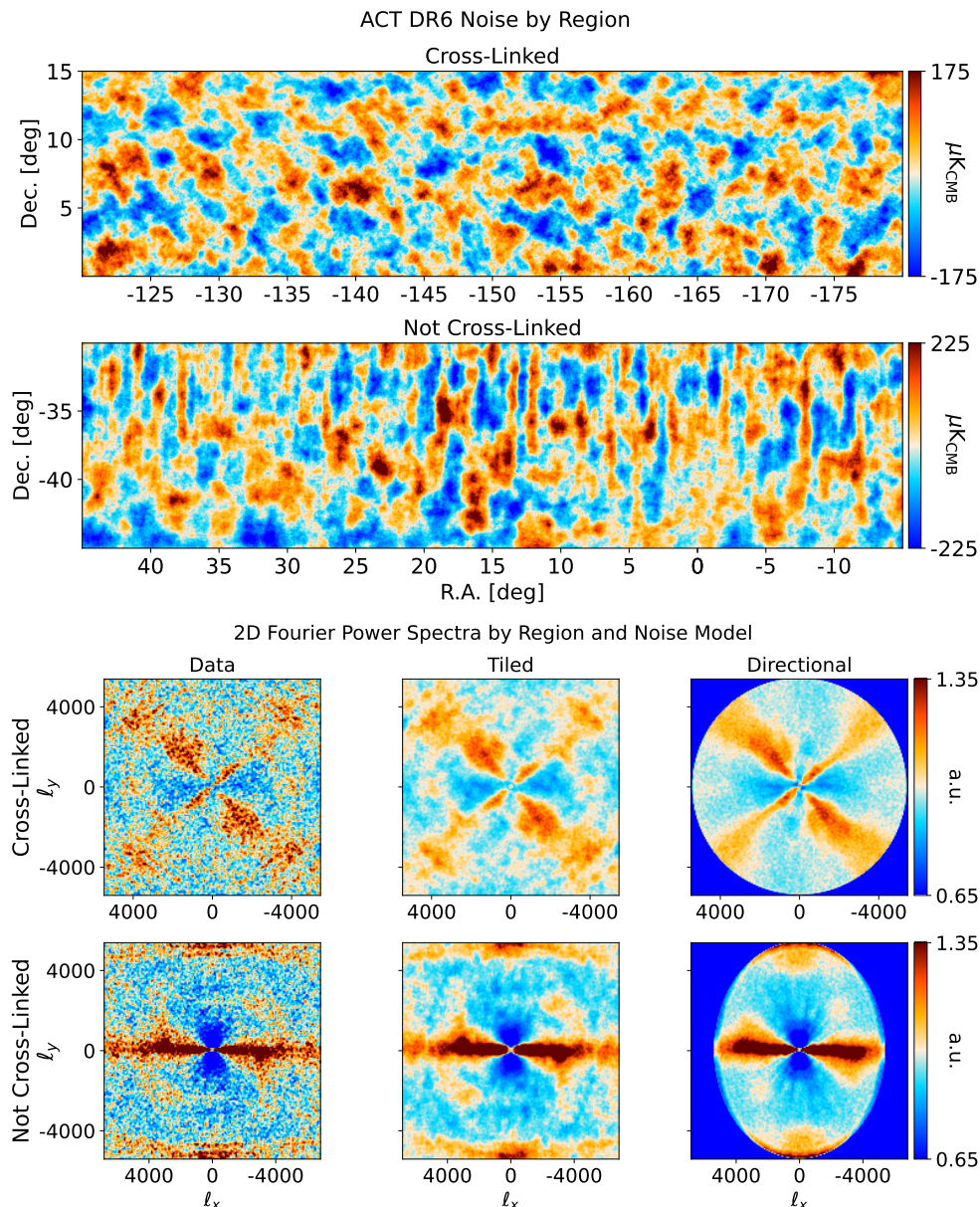
We model the noise at the map level (as measured by differences of map splits) following ref. [7], which describes two models that capture the most relevant features: a “tiled” and a “directional wavelet” noise model. The tiled model subdivides the map into a set of interleaved patches (“tiles”) and builds the noise covariance in each tile’s 2D Fourier space. The wavelet model instead subdivides the Fourier transform of the map into interleaved patches (“wavelets”) and builds the noise covariance in each wavelet’s map space. These noise models are not necessarily complete, but simulations drawn from them are sufficiently accurate for correcting and evaluating the covariance matrix. Figure 2 demonstrates an

<sup>5</sup>When discussing a field throughout the paper, we take “inhomogeneous” to mean “spatially-dependent variance over 2D positions in a map” and “anisotropic” to mean “azimuthally-dependent covariance when at a fixed 2D position in a map.”

<sup>6</sup>To streamline section 3, we do not discuss other effects in the data, such as calibration and pixel window functions, here. Those are detailed in appendix A.

<sup>7</sup>The matrix exponent ensures that  $\mathbf{S}^{\frac{1}{2}}\boldsymbol{\eta}_s^h$  is a *sample* of spherical harmonic coefficients whose covariance is the power spectrum.

<sup>8</sup>We do not measure significant beam anisotropy in DR6 and do not account for any in this paper.



**Figure 2.** *First and second rows:* the noise in the first temperature split map for PA5 f090 measured in a  $900 \text{ deg}^2$  well-cross-linked region of the ACT scan strategy. The second row shows a region with less cross-linking where scans only move in the vertical (Dec.-only) direction. *Third and fourth rows, left:* 2D Fourier noise power spectra of the first temperature split map for PA5 f090. The average radial profiles of the power spectra have been divided-out to better highlight their anisotropic patterns. The third row is measured in the well-cross-linked region, where crossing scans induce the “x-like” bars in the power spectrum. The fourth row is measured in the poorly-cross-linked region, where the Dec.-aligned scans induce vertical noise stripes in the maps that appear as a horizontal bar in the power spectrum. *Center and right:* 2D power spectra from noise simulations following ref. [7] (drawn from the tiled and directional wavelet models). The tiled model has a bandlimit of  $\ell_{\text{max}} = 10,800$ ; the wavelet model has  $\ell_{\text{max}} = 5,400$ , visible as a hard edge in 2D Fourier space. Adapted from [7]. The Author(s). CC BY 4.0.

example of the efficacy of each model at reproducing the spatially-varying noise stripiness in the data. We are agnostic as to which model performs best; rather, their existence allows us to assess robustness of the covariance matrix to the assumed noise model. Time-domain-based simulations, a possible alternative, are too computationally expensive to be practical.

As described in [7] and appendix C, samples drawn from either model follow similar prescriptions:

$$\begin{aligned}
 \mathbf{n}_i &= \boldsymbol{\sigma}_i \mathbf{Y} \mathbf{N}_i^{\frac{1}{2}} \mathbf{Y}^\dagger \boldsymbol{\Omega} (\mathcal{T}_b^\dagger \mathbf{F}^\dagger \mathcal{N}_i^{\frac{1}{2}}) \boldsymbol{\eta}_{n,i}^{\mathcal{T}f} & (\text{Tiled}) \\
 \mathbf{n}_i &= \boldsymbol{\sigma}_i \mathbf{Y} \mathbf{N}_i^{\frac{1}{2}} \mathbf{Y}^\dagger \boldsymbol{\Omega} (\mathbf{F}^\dagger \mathcal{T}_b^\dagger \mathbf{F} \mathcal{N}_i^{\frac{1}{2}}) \boldsymbol{\eta}_{n,i}^{\mathcal{T}w}, & (\text{Dir. Wavelet})
 \end{aligned} \tag{3.3}$$

where, again, we have suppressed the array ( $I$ ) and polarization ( $X$ ) labels for simplicity. Reading from right to left,  $\boldsymbol{\eta}_n^{\mathcal{T}\{f,w\}}$  is a white-noise vector in either tiled 2D Fourier space or wavelet map space,  $\mathcal{N}$  is the noise covariance in the same space,  $\mathbf{F}$  is the unitary Discrete Fourier Transform (DFT), and  $\mathcal{T}_b$  is the “backward tiling transform” that reverts the map or Fourier space subdivision (for the tiled model or directional wavelet model, respectively; see ref. [7] for more technical detail). The terms in parentheses are the defining feature of either noise model: they are what introduce the spatially-varying noise anisotropy. The sample then undergoes spherical harmonic “analysis” (i.e., from map to harmonic space), denoted by  $\mathbf{Y}^\dagger \boldsymbol{\Omega}$ , where  $\boldsymbol{\Omega}$  is a diagonal matrix in map space containing spherical harmonics analysis quadrature weights that approximately equal the area of each pixel. Finally,  $\mathbf{N}$  is the estimated noise power spectrum, and  $\boldsymbol{\sigma}$  contains the per-pixel noise standard deviations from section 2.<sup>9</sup> The noise realization,  $\mathbf{n}$ , as well as components of the noise model, depend on the map split  $i$ . We emphasize that a realistic model of the ACT map noise, compared to the signal model in equation (3.2), is inhomogeneous and anisotropic.

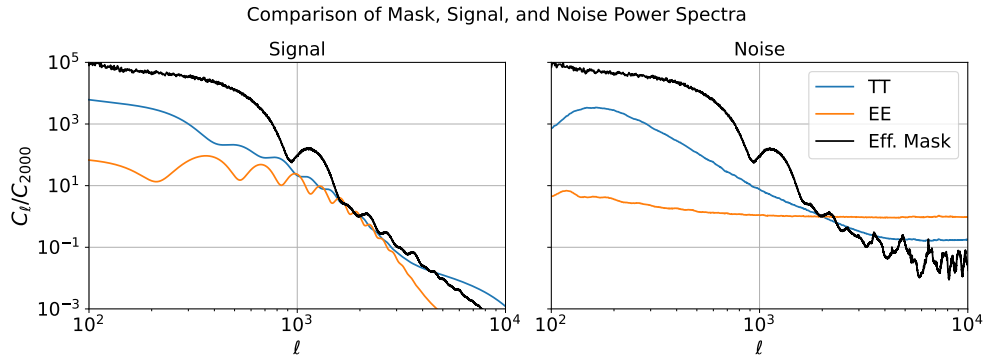
Two steps in the ACT DR6 analysis that affect the data model are the application of a Fourier-space filter and an apodized analysis mask to the data maps. Fourier modes are removed from the maps for which  $|\ell_x| \leq 90$  and  $|\ell_y| \leq 50$ ; this is the same filter as in ACT DR3 [54] and DR4 [16]. This filter removes pickup at large scales in the horizontal (R.A.-aligned) direction, and noise-dominated modes at large scales in the vertical (Dec.-aligned) direction. This filter introduces anisotropy to the signal component of the maps, while adding to the preexisting map noise anisotropy. The analysis masks described in section 2 are applied after the Fourier-space filter.

Combining these processes with equations (3.1), (3.2), and (3.3) yields the following data model for the masked maps (e.g., in the case of the “tiled” noise model):

$$\begin{aligned}
 \tilde{\mathbf{m}}_i &\equiv \mathbf{W} \mathbf{F}^\dagger \mathbf{X}_f \mathbf{F} \mathbf{m}_i = \mathbf{W} \mathbf{F}^\dagger \mathbf{X}_f \mathbf{F} \mathbf{Y} \mathbf{B} \mathbf{S}^{\frac{1}{2}} \boldsymbol{\eta}_s^h \\
 &\quad + \mathbf{W} \mathbf{F}^\dagger \mathbf{X}_f \mathbf{F} \boldsymbol{\sigma}_i \mathbf{Y} \mathbf{N}_i^{\frac{1}{2}} \mathbf{Y}^\dagger \boldsymbol{\Omega} \mathcal{T}^\dagger \mathbf{F}^\dagger \mathcal{N}_i^{\frac{1}{2}} \boldsymbol{\eta}_{n,i}^{\mathcal{T}f},
 \end{aligned} \tag{3.4}$$

where  $\mathbf{X}_f$  denotes the Fourier-space filter (which is diagonal in Fourier space), and  $\mathbf{W}$  denotes the analysis mask (which is diagonal in map space). While realistic, this inhomogeneous and anisotropic data model is not amenable to the MASTER framework: as we describe in section 3.2, the MASTER framework assumes a data model consisting of an analysis mask,  $\mathbf{W}$ , and otherwise statistically homogeneous and isotropic fields.

<sup>9</sup>The noise power spectrum,  $\mathbf{N}$ , is estimated after normalizing the maps by  $\boldsymbol{\sigma}$ .



**Figure 3.** Signal and noise power spectra compared to the power spectra for their effective masks (for PA6 f150, first split). The difference between the signal and noise effective masks is shown in figure 1: the effective mask for the noise includes the inhomogeneous survey depth and so has more structure than the signal mask, as reflected in its wider mask power spectrum. For both the signal and noise, the power spectrum of the mask does not appear to be significantly more compact, or steep, than the power spectrum of the field itself, calling the NKA into question for ACT. All power spectra are normalized at  $\ell = 2,000$ .

### 3.2 Covariance matrices in MASTER

The MASTER method includes the computation of both a power spectrum estimator,  $\hat{C}_\ell$ , and a pseudospectrum covariance matrix,  $\tilde{\Sigma}_{\ell\ell'}$ . For a data model consisting of an isotropic, scalar Gaussian field,  $a$ , sampled from the power spectrum  $C_\ell$ , and masked by an analysis window  $w$  at sky positions  $x$  with

$$\tilde{a}(x) \equiv w(x)a(x), \quad (3.5)$$

the MASTER power spectrum estimator is given<sup>10</sup> by:

$$\begin{aligned} \hat{C}_\ell &= \sum_{\ell_1} M_{\ell\ell_1}^{-1} \hat{C}_{\ell_1} \\ M_{\ell\ell_1} &\equiv (2\ell + 1) \Xi_{\ell\ell_1}(w, w) \\ \hat{C}_\ell &\equiv \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m} \tilde{a}_{\ell m}^*, \end{aligned} \quad (3.6)$$

[e.g., 18, 24, 30], where  $M_{\ell\ell_1}$  is the mode-coupling matrix,  $\Xi_{\ell\ell_1}$  is the symmetric coupling matrix [24], which we refer to as the “coupling,”  $\hat{C}_\ell$  is the *pseudospectrum* estimate, and  $\tilde{a}_{\ell m}$  is the SHT of  $\tilde{a}(x)$ . The coupling is only a function of the analysis mask via its power spectrum [39],  $W_\ell \equiv 1/(2\ell + 1) \sum_m w_{\ell m} w_{\ell m}^*$ . Under the data model of equation (3.5), the power spectrum estimator in equation (3.6) is exact and calculable in  $\mathcal{O}(\ell_{\max}^3)$ -time.

Unfortunately, there is no known analytical expression for the covariance of the pseudospectrum estimator that is also calculable in  $\mathcal{O}(\ell_{\max}^3)$ -time. A naive computation leads to an expression whose complexity is  $\mathcal{O}(\ell_{\max}^6)$  [14]. Recently, ref. [12] derived an algorithm

<sup>10</sup>This paper, and the standard MASTER framework, assume that the window  $w$  is not correlated with underlying field. Recent work by ref. [68] has derived modifications to the framework when this assumption is invalid.

with  $\mathcal{O}(\ell_{\max}^5)$  complexity. Instead, many analyses [see e.g., 22, 31, 62, 63] have employed a pseudospectrum covariance matrix with the following form:

$$\tilde{\Sigma}_{\ell\ell'} \equiv \langle (\hat{C}_\ell - \langle \hat{C}_\ell \rangle)(\hat{C}_{\ell'} - \langle \hat{C}_{\ell'} \rangle) \rangle \approx 2C_{(\ell,\ell')}^2 \Xi_{\ell\ell'}(w^2, w^2), \quad (3.7)$$

where the  $(\ell, \ell')$  subscript denotes some symmetric function of  $C_\ell$  and  $C_{\ell'}$ . As it uses the coupling matrix from equation (3.6), this expression is also calculable in  $\mathcal{O}(\ell_{\max}^3)$ -time; however, it requires invoking the ‘‘narrow kernel approximation’’ [NKA, see 30], which states that the power spectrum of the mask,  $W_\ell$ , is significantly more compact in  $\ell$  than that of the underlying field. For the mask, signal, and noise power spectra in ACT, it is not clear whether this approximation is valid (see figure 3).

Furthermore, regardless of the NKA, the MASTER formalism requires a simple data model like that in equation (3.5). Expressed in the symbolic formalism of section 3.1, this model looks like:

$$\mathbf{a} = \mathbf{WY}\mathbf{C}^{\frac{1}{2}}\boldsymbol{\eta}^h, \quad (3.8)$$

where  $\mathbf{C}$  is a diagonal matrix whose elements are given by  $C_\ell$ . As defined in equation (3.4), ACT’s signal field differs from this simpler model due to the Fourier-space filter, while the noise field also differs due to its intrinsically inhomogeneous and anisotropic structure. Thus, in order to work with MASTER methods, we need an approximate data model whose signal and noise better resemble equation (3.8), at the cost of reduced realism.

### 3.3 Approximate data model

Our covariance matrix can only involve analysis masks and isotropic power spectra as inputs (as in equation (3.8)). Here, we describe the approximations to equation (3.4) that enable this.

Because the MASTER data model does not permit intrinsic inhomogeneity nor anisotropy, we construct an isotropic and homogeneous expression for the noise by eliminating the anisotropic operations from equation (3.4). Doing so for either noise model yields:

$$\mathbf{n}_i = \sigma_i \mathbf{Y}\mathbf{N}_i^{\frac{1}{2}} \mathbf{Y}^\dagger \boldsymbol{\Omega} \boldsymbol{\eta}_{n,i}^m, \quad (3.9)$$

where  $\boldsymbol{\eta}_n^m$  is a white-noise vector in map space. As we show in appendix A, this noise model is approximately equal to the following expression:

$$\mathbf{n}_i = \sigma_i \boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{Y}\mathbf{N}_i^{\frac{1}{2}} \boldsymbol{\eta}_{n,i}^h \quad (\text{Approximate}) \quad (3.10)$$

where, again,  $\boldsymbol{\eta}_n^h$  is instead a white-noise vector in harmonic space. This noise model states that we draw a Gaussian realization from an isotropic noise power spectrum, and then weight each pixel by the square-root of its area and its noise standard deviation. While incurring a ‘‘data model error,’’ its ingredients are now identical in form to those of the signal, where  $\sigma_i \boldsymbol{\Omega}^{\frac{1}{2}}$  is an effective noise-weight mask.

The Fourier-space filter also introduces anisotropy to both the signal and noise. Unlike the intrinsic noise anisotropies, the anisotropy due to the filter is too large to be neglected and must be accounted for analytically. The core of the approximation is treating the filter in

harmonic space, rather than Fourier space. We give further detail on the exact approximation we use in section 4 and appendix B.2; for the purposes of this section, we absorb the filter into the definition of the signal and noise power spectra as an isotropic transfer function,  $\mathbf{T}^\alpha$ , in equations (3.6) and (3.7), where  $\alpha$  is an exponent that helps control the function shape, and  $\mathbf{T}$  is diagonal over  $\ell$  (i.e., its diagonal is  $t_\ell^\alpha$ ). While similar to ref. [16], the treatment of  $\mathbf{T}$  in this paper does not assume that it enters those equations with fixed exponents, but rather we find these exponents are functions of other inputs, predominantly the filter itself.

Taken together, we have the following approximate data model for the ACT DR6 maps:

$$\tilde{\mathbf{m}}_i = \mathbf{W}\mathbf{Y}\mathbf{B}(\mathbf{T}^\alpha\mathbf{S})^{\frac{1}{2}}\boldsymbol{\eta}_s^h + \mathbf{W}\boldsymbol{\sigma}_i\boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{Y}(\mathbf{T}^\alpha\mathbf{N}_i)^{\frac{1}{2}}\boldsymbol{\eta}_{n,i}^h. \quad (3.11)$$

This simplified model is compatible with the MASTER formalism in analogy with equation (3.8). To be clear, we do not expect equation (3.11) to correctly model the anisotropic ACT data at the map level. Rather, it is an attempt to preserve as much of the data realism as possible while still permitting the use of the MASTER covariance framework.

As discussed in the next section, the form of the covariance matrix that results from equation (3.11) closely resembles that in ref. [62]. Unlike for *Planck*, however, it requires assuming the approximate ACT data model is sufficiently accurate at the level of the power spectrum covariance matrix. We investigate the validity of both the NKA and the approximate data model in section 5.

We emphasize that the need for such an investigation is not abstract. Because equation (3.11) is incorrect for the ACT data, the covariance estimate derived from it in the next section is *biased*. This bias is analogous to the “noise bias” occurring in power spectrum estimates from maps with correlated noise. Unlike the noise bias for the power spectrum estimator, which can be circumvented through the use of cross-spectra [64, 70], this “covariance bias” is unavoidable and its size must be checked empirically.

## 4 Covariance matrix pipeline

In this section, we describe how we construct the inhomogeneous ACT DR6 power spectrum covariance matrix. First, we assemble the analytic part, assuming the NKA and the approximate data model in equation (3.11) (section 4.1). In parallel, we construct an ensemble of simulations following the ACT DR6 data model of equation (3.4) (section 4.2). Lastly, we use the Monte Carlo covariance matrix formed by the simulations to correct approximation-induced errors in the analytic matrix (section 4.3). This matrix is only the disconnected, Gaussian part of the covariance; beam measurement uncertainty, and the subdominant non-Gaussian contributions to the covariance from CMB lensing, clusters, and point sources will be discussed in the ACT DR6 power spectrum paper.

### 4.1 Analytic covariance matrix

We generalize the MASTER pseudospectrum covariance in equation (3.7) to the case of multiple fields and datasets. We write equation (3.11) in the form of equation (3.5), including all metadata labels:

$$\begin{aligned} m_{I_i}^X(x) &= w_I^X(x)s_I^X(x) + w_I^X(x)\sigma_{I_i}^X(x)\boldsymbol{\Omega}^{\frac{1}{2}}(x)n_{I_i}^X(x) \\ &\equiv u_I^X(x)s_I^X(x) + v_{I_i}^X(x)n_{I_i}^X(x), \end{aligned} \quad (4.1)$$

where we have defined the effective signal and noise masks ( $u$  and  $v$ ),  $s_I^X(x)$  is the realization from the isotropic signal power spectrum including the beam ( $S_{IJ,\ell}^{XY}$ ), and  $n_{I_i}^X(x)$  is the realization from the isotropic noise power spectrum ( $N_{I_i J_j, \ell}^{XY}$ ). We have assumed that the analysis masks ( $w$ ) and signal realizations do not depend on the split ( $i$ ). Thus, the map pseudospectra carry a pair of metadata labels, and the covariance carries two pairs.

Assuming the signal and noise are uncorrelated, we obtain for the pseudospectrum covariance matrix:

$$\begin{aligned} \tilde{\Sigma}_{I_i J_j, P_p Q_q, \ell \ell'}^{WX,YZ} &= S_{IP,(\ell,\ell')}^{WY} S_{JQ,(\ell,\ell')}^{XZ} \Xi_{\ell \ell'}^{\beta(WY,XZ)}(u_I^W u_P^Y, u_J^X u_Q^Z) \\ &+ S_{IP,(\ell,\ell')}^{WY} N_{JQ,(\ell,\ell')}^{XZ} \Xi_{\ell \ell'}^{\beta(WY,XZ)}(u_I^W u_P^Y, v_{J_j}^X v_{Q_q}^Z) \\ &+ N_{I_i P_p,(\ell,\ell')}^{WY} S_{JQ,(\ell,\ell')}^{XZ} \Xi_{\ell \ell'}^{\beta(WY,XZ)}(v_{I_i}^W v_{P_p}^Y, u_J^X u_Q^Z) \\ &+ N_{I_i P_p,(\ell,\ell')}^{WY} N_{J_j Q_q,(\ell,\ell')}^{XZ} \Xi_{\ell \ell'}^{\beta(WY,XZ)}(v_{I_i}^W v_{P_p}^Y, v_{J_j}^X v_{Q_q}^Z) \\ &+ (Y, P_p) \leftrightarrow (Z, Q_q), \end{aligned} \quad (4.2)$$

where we introduce the coupling “spin”,  $\beta$ , which can be one of “00”, “0+”, “++”, or “--” [11]. The dependence of  $\beta$  on the four input polarizations is given in appendix B.1. The argument of the coupling denotes the cross-power spectrum of a pair of “masks” that are each a product of two effective signal or noise masks. The arrow indicates that the rest of the expression is formed by the interchange of the third and fourth fields. We derive this result in appendix B.1, although it is also presented implicitly in ref. [62] and ref. [59]. The analogous expression for a single field (e.g., the “signal-only” covariance), or for the special case that the effective signal and noise masks are all equivalent ( $u = v$ ), is a standard result in the literature [see e.g., 11, 18, 25, 30, and also our discussion in section 4.1.6]. Besides the coupling spin, there is no fundamental difference between equations (3.7) and (4.2) — the complication is just in the (expansive) bookkeeping of signal, noise, array, polarization, and split.

We use a modified version of the traditional NKA — the “improved” NKA [INKA, 59] — in light of the steepness of the CMB and ACT noise spectra in figure 3. Here, each power spectrum in equation (4.2) is replaced by a normalized pseudospectrum:

$$C_{IJ,\ell}^{XY} \rightarrow \tilde{C}_{IJ,\ell}^{XY}/w_2(w_I^X, w_J^Y), \quad (4.3)$$

where  $C$  is either  $S$  or  $N$ ,  $w_I^X$  and  $w_J^Y$  are the effective masks for arrays (polarizations)  $I$  and  $J$  ( $X$  and  $Y$ ), and  $w_2(a, b) = \sum_x a(x)b(x)\Omega(x)/4\pi$ . The pseudospectra are related to the power spectra via the MASTER mode-coupling matrices: the expressions are a standard result in the literature [e.g., 3, 44]. Generalizing equation (3.6), the mode-coupling matrices are a function of the cross-power spectrum of  $w_I^X$  and  $w_J^Y$ . Ref. [59] found the INKA improves the accuracy of analytic MASTER covariance matrices when applied to galaxy weak lensing, and argued for its superiority in all cases. It has seen increasing uptake in (for example) refs. [22, 27, 34, 71], but to our knowledge has not yet been used in a CMB analysis.

As in ref. [59], we choose to use the arithmetic mean to define the symmetric function  $C_{(\ell,\ell')} \equiv (C_\ell + C_{\ell'})/2$ , where, again,  $C$  is either  $S$  or  $N$ . Ref. [29] argues that this choice is best motivated under the NKA, and also has the advantage of avoiding numerical issues for cross-spectra that can be negative (e.g., as is the case for the geometric mean and the  $TE$  spectrum). We combine this with equation (4.3) when substituting spectra into equation (4.2).

The preceding prescription defines the following required inputs:

- Couplings for each spin and combination of four effective masks (for the pseudospectrum covariance matrix).
- Couplings for each spin and combination of two effective masks (for the INKA mode-coupling matrices).
- $w_2$  factors for each combination of two effective masks (for the INKA).
- Fiducial signal and noise power spectra.

We briefly elaborate on these elements in the following.

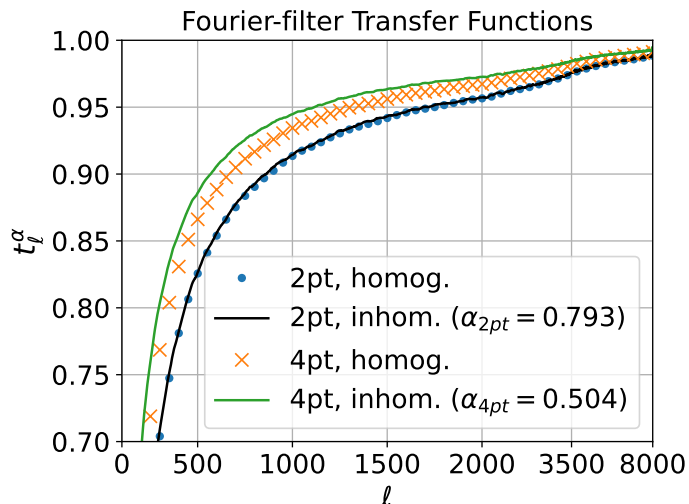
#### 4.1.1 Coupling matrices

We calculate all unique couplings required for the pseudospectrum covariance, as described above. While the combinatorics for the two-mask couplings are manageable, they become prohibitive for the four-mask couplings if computed naively, with almost 2 million possible combinations. To conserve resources, we tabulate the number of unique couplings (e.g., given that the DR6 analysis masks do not depend on polarization). Then, we discard couplings that would correspond to a cross-split noise power spectrum in equation (4.2): since we assume splits have independent noise, these terms would be zero regardless of the coupling. This results in only 3,062 couplings to compute. Finally, we accelerate computation by using the Toeplitz approximation with the same default parameters as ref. [53]. We find errors incurred by this approximation are sub-percent. In parallel, we calculate all required  $w_2$  factors at negligible cost.

#### 4.1.2 Fourier-space filter treatment

As discussed in section 3.3, we model the  $\mathcal{O}(1)$  effect of the Fourier-space filter on the power spectrum covariance. We do this by applying an isotropic transfer function,  $t_\ell^\alpha$ , to the fiducial signal and noise power spectra (equation (3.11)). The filter has a different effect depending on whether the fiducial spectrum appears in the power spectrum estimator or the pseudospectrum covariance matrix, which we model by an isotropic transfer function with different shapes. We capture the different shapes by modifying the exponent of the transfer function,  $\alpha$ . In other words, when the filtered power spectrum appears in the MASTER pseudospectrum estimator, we multiply it by the “two-point” transfer function given by  $t_\ell^{\alpha_{2pt}}$ . When the filtered power spectrum appears as part of the NKA (or INKA) in the covariance matrix, we multiply it by the “four-point” transfer function,  $t_\ell^{\alpha_{4pt}}$ . In this section, we summarize how we determine each component of this model:  $t_\ell$ ,  $\alpha_{2pt}$ , and  $\alpha_{4pt}$ . A detailed description of each step is given in appendix B.2.

We first obtain the transfer function template,  $t_\ell$ , by applying the Fourier filter to a small number (50) of full-sky, white-noise simulations. Because the power spectrum for this field is a known constant, any observed deviations in the power spectra of the simulations are due to the filter, independent of any masking. We apply a fourth-order Savitzky-Golay filter to the simulated spectra to construct a smooth template.



**Figure 4.** The “two-point” and “four-point” Fourier-space filter transfer functions have different shapes. Comparing our new method as part of the inhomogeneous matrix to the homogeneous matrix approach, the two methods agree on the two-point transfer functions to  $< 1\%$  but disagree on the four-point transfer functions on medium and large scales. The homogeneous matrix transfer function is determined by bin rather than by  $\ell$ . This figure shows the temperature case.

Next, we fit for the two-point ( $\alpha_{2pt}$ ) and four-point ( $\alpha_{4pt}$ ) shape exponent. We do so by running a larger number (500) of simulations, which are realizations from a mock, noise-like power spectrum (one for temperature and one for polarization). These simulations are filtered, and their pseudospectra measured using a representative, average effective mask. In the two-point case, we fit for  $\alpha_{2pt}$  by forward modeling the observed pseudospectra as  $\tilde{C}_\ell(\alpha_{2pt}) = \sum_{\ell_1} M_{\ell\ell_1} t_{\ell_1}^{\alpha_{2pt}} C_{\ell_1}$ , and analogously for  $\alpha_{4pt}$  by forward modeling  $\tilde{C}_\ell(\alpha_{4pt})$  in an expression for the power spectrum covariance matrix (see appendix B.2). Our choice to use a mock power spectrum, and single representative effective mask, is justified by our finding that both  $\alpha_{2pt}$  and  $\alpha_{4pt}$  are largely independent of the mask and power spectrum; instead, they are mainly a function of the actual Fourier-filter. Thus, we can reuse the model for both signal and noise, as well as across arrays and splits. This feature drastically limits the computational cost of these dedicated simulations in comparison to the full-dataset ensemble in section 4.2. Results of this process are shown in figure 4 for the temperature case.

The ACT DR4 analysis [16] also used simulations to determine the approximate two-point transfer function from the Fourier-filter, although the specific implementation differed from this paper. However, as roughly mapped onto our method, ref. [16] used the fixed relation  $\alpha_{4pt} = 0.75\alpha_{2pt}$ . While a reasonable approximation for our specific Fourier-filter, we find this relation does not generally hold. A comparison of our new model to that of ref. [16] (which is also used in the homogeneous matrix) is shown in figure 4, where the difference can be seen on large scales for the four-point transfer function (note,  $0.75\alpha_{2pt} = 0.595$ , whereas we find  $\alpha_{4pt} = 0.504$ ). Because the covariance matrix is quadratic in the power spectra, our new method increases the analytic covariance of each spectrum by  $\sim 5\%$  at  $\ell = 500$ . In either case, inaccuracies in the Fourier-filter treatment are corrected using the Monte Carlo covariance matrix discussed in section 4.2.

Finally, we note that our procedure for fitting  $\alpha_{2pt}$  and  $\alpha_{4pt}$  does not *exclusively* capture the effect of the Fourier-space filter. Rather, it can capture *any* deviations from the expected covariance under the MASTER framework. Thus, while targeting the anisotropic Fourier-filter, the model may also capture deviations due to a breakdown of the NKA. It is important to bear this in mind when we evaluate the performance of the “analytic” covariance matrix: any purely NKA-induced errors may be suppressed by the Fourier-filter correction.

### 4.1.3 Fiducial signal power spectra

We use a fiducial signal power spectrum including the CMB, extragalactic foregrounds, and dust, following the cosmology and foreground model of ref. [16]. Thus, the signal cross-power spectrum of a given pair of arrays is a function of the array passbands (which we take to be fixed) and the spectral energy distributions (SEDs) of each component. The parameters of the model used in this paper are given in table 3 in appendix B.3. Following refs. [13, 63], we assume the fiducial signal power spectra are sufficiently close to the true power spectra that the ACT DR6 cosmological results are unbiased. We then apply the beam transfer functions for each array, and, following section 4.1.2, we account for the four-point effect of the Fourier-space filter by applying  $t_\ell^{\alpha_{4pt}}$ . Finally, following the INKA, we convert the power spectra into pseudospectra using the mode-coupling matrix for the two arrays’ effective signal masks.

### 4.1.4 Fiducial noise power spectra

We measure the noise spectra directly from the data. Following ref. [7], we assume uncorrelated noise between maps from different physical detector wafers, but possibly correlated noise between frequency channels on the same wafer. For example, we take PA5 f090 noise to be independent from PA6 f090 noise, but not from PA5 f150 noise. We perform the following steps for each detector wafer, and for each frequency-polarization pair on that wafer. First, we compute the average pseudospectrum over pairs of different map splits (for four splits, there are 12 such pairs).<sup>11</sup> Because the ACT DR6 map splits contain independent noise, this average represents an unbiased signal-only model. We then compute the pseudospectrum for each map split paired with itself (i.e., an “auto-split” pseudospectrum). This second quantity contains the same signal as the first average in expectation, but with an additive noise bias equal to the noise power spectrum in that split. Subtracting the signal-only model from the “auto-split” pseudospectrum results in a measurement of the noise-only model. We then reduce statistical scatter by filtering the measured noise pseudospectra with a fourth-order Savitzky-Golay filter.

Because the data from which we measure the noise pseudospectra have been filtered with the Fourier-space filter, they represent mode-coupled noise power spectra after applying the *two-point* Fourier-filter transfer function. However, following section 4.1.2, the covariance matrix recipe calls for mode-coupled noise power spectra after applying the *four-point* Fourier-filter transfer function. Thus, we first apply the inverse of the mode-coupling matrix for the two arrays’ and splits’ effective noise masks and then divide-out  $t_\ell^{\alpha_{2pt}}$  to get “unfiltered”

<sup>11</sup>The average over 12 pairs is correct regardless of whether the frequency-polarization legs in the pair are different or the same. In the latter case, it is true that we double computation, but our code is simpler and easier to read.

noise power spectra. Then, like the signal spectra, we apply  $t_\ell^{\alpha_{4pt}}$  and convert back into noise pseudospectra using the same mode-coupling matrix.

#### 4.1.5 Pseudospectrum to power spectrum covariance

The preceding ingredients produce pseudospectrum covariance matrix blocks at the array, split, and polarization level via equation (4.2). To convert to the power spectrum covariance matrix, we apply the matrix operation that transforms from pseudospectra to power spectra. For a given measured pair of arrays ( $IJ$ ) and polarizations ( $XY$ ), the power spectrum estimator (in symbolic form) is given by:

$$\begin{aligned}\hat{\mathbf{D}}_{IJ}^{XY} &= \mathbf{U}_{IJ}^{XY} \mathbf{P} \mathbf{M}_{IJ}^{-1,XY} \frac{1}{N_{\{ij\}}} \sum_{\{ij\}} \hat{\mathbf{C}}_{I_i J_j}^{XY} \\ &\equiv \mathbf{Q}_{IJ}^{XY} \sum_{\{ij\}} \hat{\mathbf{C}}_{I_i J_j}^{XY}.\end{aligned}\tag{4.4}$$

Reading from right to left,  $\hat{\mathbf{C}}_{I_i J_j}^{XY}$  is the vector over  $\ell$  of the pseudospectrum for the split-pair  $ij$ , and the sum is over the  $N_{\{ij\}} = 12$  split-pairs for which  $i \neq j$ . As in section 4.1.4, this aspect of the power spectrum pipeline ensures the estimator contains no noise bias, regardless of the noise model [64, 70]. The matrix  $\mathbf{Q}_{IJ}^{XY}$  is composed<sup>12</sup> of the inverse mode-coupling matrix,  $\mathbf{M}_{IJ}^{-1,XY}$ , a “weighting-binning” matrix,  $\mathbf{P}$ , which transforms  $C_\ell \rightarrow D_\ell$ , with  $D_\ell \equiv \ell(\ell+1)/(2\pi)C_\ell$  and bins the resulting  $\mathbf{D}$  vector over  $\ell$ , and  $\mathbf{U}_{IJ}^{XY}$ , which is the power spectrum pipeline’s implementation of the two-point Fourier-filter correction. This correction is more precise than that described in section 4.1.2: it is based on more simulations and occurs at the binned power spectrum level, as opposed to the unbinned power spectrum level. We calculate  $\mathbf{Q}$  explicitly. Because the covariance is bilinear, we have for a block of the binned power spectrum covariance matrix,  $\Sigma_{bb'}$ :

$$\Sigma_{IJ,PQ,bb'}^{WX,YZ} = \sum_{\ell} Q_{b\ell} \sum_{\ell'} Q_{b'\ell'} \sum_{\{ij\}} \sum_{\{pq\}} \tilde{\Sigma}_{I_i J_j, P_p Q_q, \ell\ell'}^{WX,YZ}.\tag{4.5}$$

Equation (4.5) defines our “analytic” covariance matrix.

For convenience, we briefly define a more-compact notation of a covariance matrix block. We combine the polarization pair  $WX$  and array pair  $IJ$  defining the block of the power spectrum vector into one index,  $\beta$ . Thus, we may also refer to any element of the binned covariance matrix as  $\Sigma_{\beta\beta',bb'}$ . We define the *block-wise* matrix diagonal to be the set of matrix blocks  $\beta' = \beta$  for any bins, and the *bin-wise* matrix diagonal to be the bins  $b' = b$  for any block; the *main* matrix diagonal is the set of elements for which  $\beta' = \beta$  and  $b' = b$ . We may also identify series of off-diagonal blocks (e.g.,  $\beta' = \beta + 1$ ) and off-diagonal bins (e.g.,  $b' = b + 1$ ) explicitly.

#### 4.1.6 Comparison to previous methods

This analytic covariance is closely related to previous analyses’ prescriptions. The analytic part of the “homogeneous” ACT matrix uses the public power spectrum covariance matrix

<sup>12</sup> $\mathbf{Q}$  nearly exactly reproduces the ACT DR6 power spectrum pipeline. It omits, however, some small, additive corrections made to the power spectrum that cannot be easily incorporated into  $\mathbf{Q}$ , see section 4.2.

code `pspy`, that like `NaMaster`<sup>13</sup> provides convenient functions that assume homogeneous noise weighting (i.e.,  $v(x) \equiv w(x)$  in equation (4.1)). Ref. [51] found that neglecting survey depth in the effective noise masks underestimates *Planck* covariance matrices by  $\sim 10\%$ , while ref. [7] found such prescriptions underestimate ACT DR6 covariance matrices by up to  $\sim 20\%$ . In addition to assuming homogeneous survey depth, the homogeneous matrix uses the standard NKA, the ACT DR4-style Fourier-space filter correction (see the discussion in section 4.1.2), a simpler fiducial noise power spectrum pipeline that averages the spectra over split-crosses and uses binning instead of an  $\ell$ -dependent filter to smooth the measurements, and it sets cross-array noise spectra to be positive and cross-polarization noise spectra to be 0. Further implementation details for the homogeneous matrix are available in the `pspy` documentation.<sup>14</sup>

The closest analytic covariance prescription to our “inhomogeneous” matrix is in ref. [62], where the inhomogeneous survey depth is accounted for in the effective noise masks, and thus the distinct signal and noise coupling terms are included as in equation (4.2). For treating the signal spectra, ref. [62] used the traditional NKA; for the noise, an “approximate” treatment of its non-white character is noted. Close inspection of ref. [61] and ref. [62] shows this treatment was equivalent to the INKA, and thus no more approximate than their signal treatment under a MASTER-like data model. Finally, that analysis uses a mix of arithmetic and geometric symmetry of spectra, unlike our use of only arithmetic symmetry, although this difference is minor. The main difference between ref. [62] and this paper are in the properties of the ACT data in comparison to *Planck*. The *Planck* noise power spectra, though not perfectly white, are broadly smooth and within  $\sim 30\%$  of white for scales smaller than  $\ell \sim 500$  (200) for temperature (polarization) [51, 62]. Due to the atmosphere, the ACT noise power spectra are steeper, and their  $1/f$  character persists to smaller scales, especially in temperature. Additionally, *Planck* noise adheres more closely to the MASTER data model, lacking the spatially-dependent stripy noise patterns that are prominent in ACT [see e.g., 58].

## 4.2 Simulations

We complement the analytic power spectrum covariance matrix of section 4.1 with a Monte Carlo covariance made from a full-dataset simulation ensemble. We generate this ensemble of simulations following the full ACT data model of equation (3.4) and appendix A. This Monte Carlo covariance serves as a noisy estimate of an unbiased covariance matrix that incorporates the full data realism.

The signal components of the simulations are drawn from the same fixed signal model as in section 4.1. The processing of the signal components includes an additive correction for  $TE$  power spectra due to the Fourier-filter that cannot be incorporated into the  $\mathbf{U}$  matrices from section 4.1.5. As a baseline, we draw the noise components from the tiled noise model (see equation (3.3)). Between the directional wavelet model and the tiled model, only the tiled model supports simulations beyond the smallest scale of the DR6 power spectrum ( $\ell_{\max} = 8,500$ ). In particular, the noise simulations are drawn at half the resolution of the data — 1 arcmin pixels — supporting scales as small as  $\ell_{\max} = 10,800$ . Further details are provided in appendix C and ref. [7].

<sup>13</sup><https://github.com/LSSTDESC/NaMaster>.

<sup>14</sup>[https://pspy.readthedocs.io/en/latest/scientific\\_doc.pdf](https://pspy.readthedocs.io/en/latest/scientific_doc.pdf).

Including signal and noise, in our baseline ensemble we draw 1,600 realizations of each array, polarization, and split. Each simulation is passed through the power spectrum reconstruction pipeline, and we use the ensemble of simulated spectra to construct the Monte Carlo covariance. We note that the simulations do not include two small corrections made to the data: beam leakage, because its uncertainty is propagated into the covariance analytically; and aberration due to Earth’s motion with respect to the CMB rest frame, because this is a second-order effect at the covariance level. This ensemble is sufficient to constrain each element of the covariance matrix to  $\lesssim 3.5\%$ ; nevertheless, drawing enough simulations to ensure a well-conditioned, non-singular Monte Carlo matrix is not feasible.

### 4.3 Simulation-based correction

The goal of the covariance pipeline is to produce an accurate, yet tractable, covariance matrix for the ACT DR6 power spectrum. While the Monte Carlo matrix is not limited by the approximations of the analytic matrix, drawing enough simulations to reach convergence is prohibitively expensive. Thus, we start by assuming the analytic matrix is a good, if slightly biased, estimate of the true covariance matrix. Then, we use the Monte Carlo matrix to correct the analytic matrix where they measurably differ. This procedure — optimizing between a biased, analytic covariance estimate and a noisy, unbiased Monte Carlo covariance estimate — is referred to as matrix “shrinkage” or “conditioning” and is a well-studied problem in applied statistics [see e.g., 48, 49, 67]. Recently, ref. [8] and ref. [52] reviewed several covariance matrix shrinkage methods for cosmology. In this paper, we opt for a new, simple shrinkage method.

Our method relies on two core assumptions: the eigenbases of the analytic and Monte Carlo covariances are close, and the ratio of their eigenspectra are smooth. Defining the matrix exponent of the covariance as:

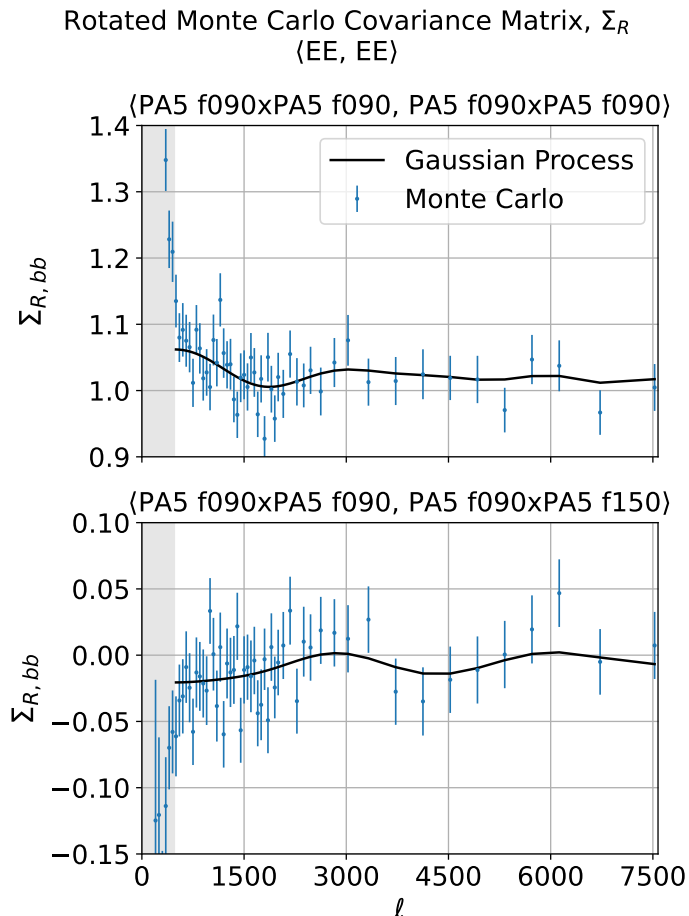
$$\boldsymbol{\Sigma}^p \equiv \mathbf{O}\mathbf{E}^p\mathbf{O}^T, \quad (4.6)$$

where  $\mathbf{O}$  is orthogonal and  $\mathbf{E}$  is diagonal, we can construct a *rotated* Monte Carlo covariance matrix:

$$\begin{aligned} \boldsymbol{\Sigma}_R &\equiv \boldsymbol{\Sigma}_A^{-\frac{1}{2}}\boldsymbol{\Sigma}_M(\boldsymbol{\Sigma}_A^{-\frac{1}{2}})^T \\ &= \mathbf{O}_A\mathbf{E}_A^{-\frac{1}{2}}(\mathbf{O}_A^T\mathbf{O}_M)\mathbf{E}_M(\mathbf{O}_M^T\mathbf{O}_A)\mathbf{E}_A^{-\frac{1}{2}}\mathbf{O}_A^T \\ &\approx \mathbf{O}_A(\mathbf{E}_M\mathbf{E}_A^{-1})\mathbf{O}_A^T, \end{aligned} \quad (4.7)$$

where  $M$  ( $A$ ) denotes the Monte Carlo (analytic) matrix, and we have used  $\mathbf{O}_A \approx \mathbf{O}_M$ . Then,  $\mathbf{E}_M\mathbf{E}_A^{-1}$  is a diagonal matrix populated by the ratio of the Monte Carlo and analytic eigenspectra. If this ratio is approximately flat, then  $\boldsymbol{\Sigma}_R$  is close to diagonal, and its elements preserve the original ordering of the data vector.

We find these approximations perform well, but are not exact. Specifically, while  $\boldsymbol{\Sigma}_R$  is diagonally-dominant with diagonal values of  $\mathcal{O}(1)$ , the bin-wise diagonals ( $b' = b$ ) of its off-diagonal blocks ( $\beta' \neq \beta$ ) are non-zero at the few-percent level. This motivates the following correction procedure: for each block (i.e., for each  $\beta\beta'$  pair), we extract the corresponding bin-wise diagonal of  $\boldsymbol{\Sigma}_R$  and smooth it using a Gaussian process with a radial basis function (RBF) kernel. The fit involves optimizing two hyper-parameters of the Gaussian process —



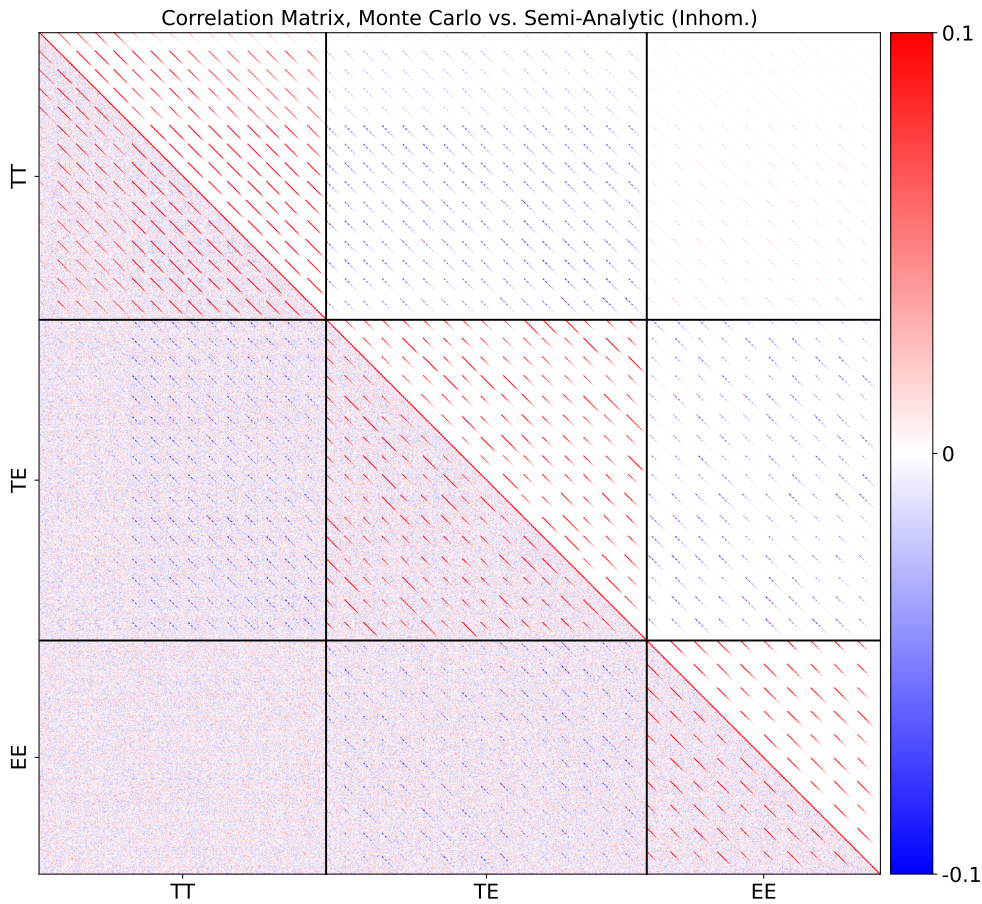
**Figure 5.** *Top:* the main diagonal of the  $\Sigma_R$  matrix defined in equation (4.7), for the PA5 f090 x PA5 f090  $EE$  block. If the analytic and Monte Carlo covariance matrices are sufficiently close, this is approximately the ratio of the eigenvalues of the two matrices. *Bottom:* the bin-wise diagonal of the covariance between the PA5 f090 x PA5 f090  $EE$  and PA5 f090 x PA5 f150  $EE$  blocks. In both cases, the  $1\sigma$  scatter of the Monte Carlo estimates are shown, with a Gaussian process fit using data above the scale-cut, indicated by the grey region. The Gaussian processes do not account for correlations between bins.

the characteristic amplitude and length scale of any coherent features in the data — given the observed data and its errors. In this case, “data” refers to the values of the  $\Sigma_R$  diagonal, and Gaussian errors on  $\Sigma_R$  are estimated directly from the scatter of the simulations. Example Gaussian process fits are shown in figure 5. We only perform the fit over  $\ell$ -bins preserved by our pre-unblinding scale cuts, discussed in section 2.<sup>15</sup> Finally, for each block, we set all bin-wise off-diagonal elements ( $b' \neq b$ ) to zero. This procedure yields the “corrected” matrix,  $\Sigma_{R,corr}$ . Finally, we rotate back to the original basis, yielding the final corrected covariance matrix:

$$\Sigma_{corr} \equiv \Sigma_A^{\frac{1}{2}} \Sigma_{R,corr} (\Sigma_A^{\frac{1}{2}})^T. \quad (4.8)$$

This procedure has no free parameters and is fast to compute.

<sup>15</sup>The noise simulations entering the Monte Carlo covariance matrix have excess power at scales larger than the scale cuts, leading to the discrepant large values at low  $\ell$  in figure 5, that would otherwise bias the fits. Also see the discussion in section 5.2.



**Figure 6.** *Below-left of main diagonal:* the correlation matrix for the Monte Carlo covariance. Most elements are zero-mean statistical noise. *Above-right of main diagonal:* the correlation matrix for the corrected analytic covariance from the inhomogeneous prescription. We apply the data and scale cuts from section 2, and restrict to the  $TT$ ,  $TE$ , and  $EE$  polarization pairs that enter the DR6 likelihood. As noted in the text, additional correlation from non-Gaussian sky components and beam uncertainty are discussed in the ACT DR6 power spectrum paper and are not plotted here.

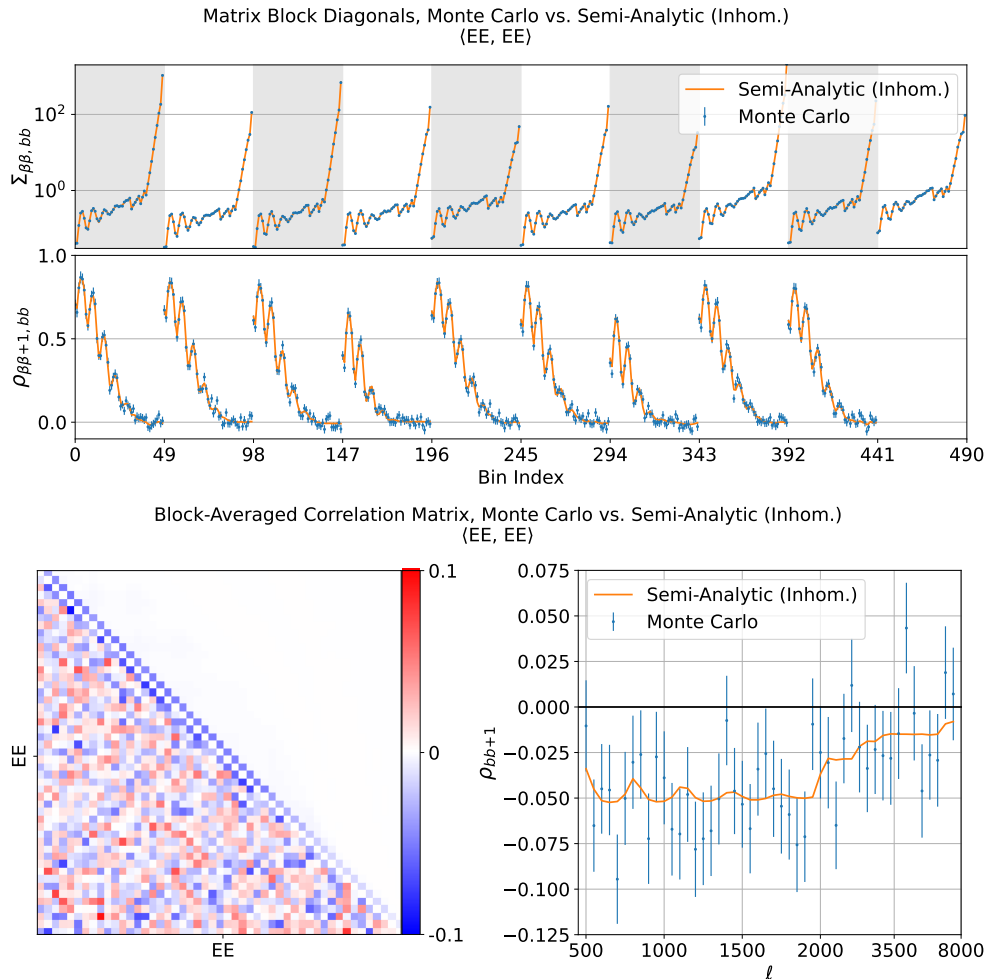
## 5 Results

Our semi-analytic covariance matrix relies on several approximations unique to the ACT DR6 data. In this section, we evaluate the accuracy of the pipeline by comparison to the Monte Carlo covariance matrix. Since we take the Monte Carlo matrix to be an unbiased estimate of the true covariance, we also probe its internal robustness.

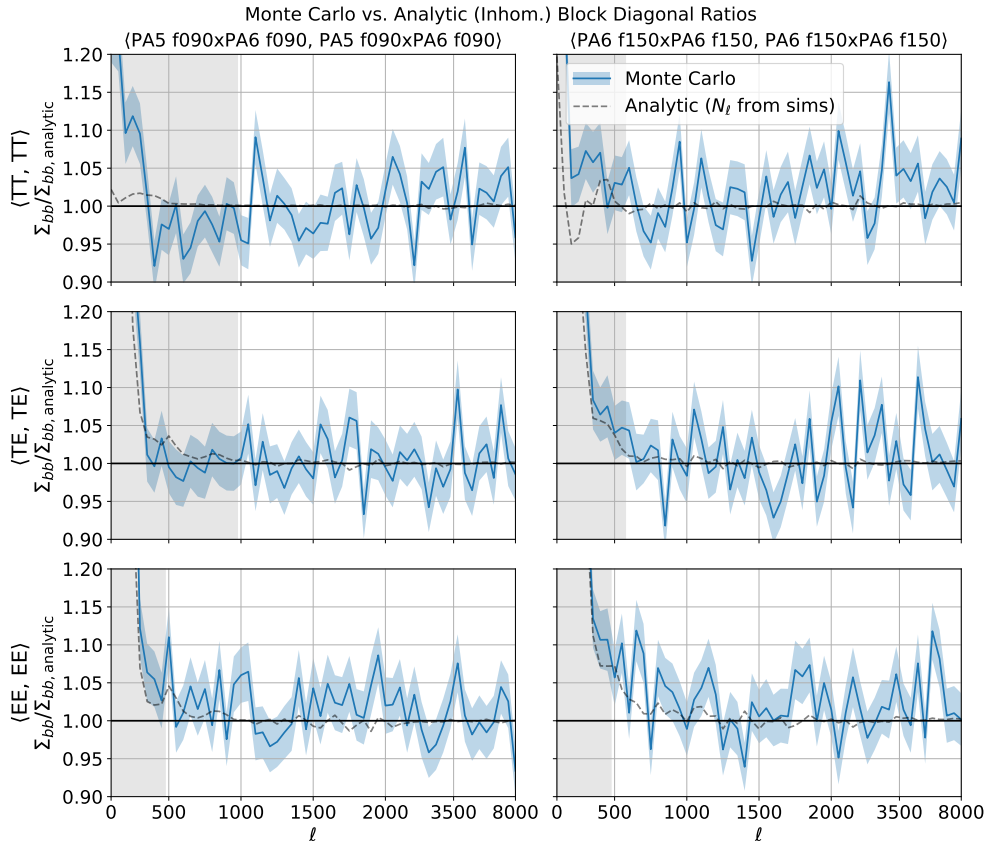
### 5.1 Covariance matrix pipeline validation

Here we show that the semi-analytic covariance matrix achieves excellent agreement with the Monte Carlo matrix, and that the size of the corrections applied in section 4.3 are small.

The structures of both the Monte Carlo matrix and the semi-analytic covariance matrix are dominated by the bin-wise diagonals ( $b' = b$ ) across all diagonal and off-diagonal blocks (all  $\beta\beta'$  pairs). We demonstrate this visually by plotting the correlation matrix of both matrices in figure 6. For a generic covariance matrix  $\Sigma_{ij}$  with elements indexed by  $ij$ , the



**Figure 7.** *First row:* main diagonal ( $\beta' = \beta, b' = b$ ) of the  $EE$  part of the covariance matrix,  $\Sigma$ . The data and scale cuts from section 2 result in 10  $EE$  array-pairs, or blocks, each containing 49 bins. The boundaries of these blocks along the main diagonal are denoted by the alternating grey and white bands. *Second row:* the bin-wise diagonal ( $b' = b$ ) of the first off-diagonal blocks ( $\beta' = \beta + 1$ ) of the  $EE$  part of the correlation matrix,  $\rho$ . There are only 9 such blocks. In both cases, the matrix contains significant structure and dynamic range. *Bottom left:* the correlation matrix after averaging over all 100 blocks of the  $EE$  part of the covariance matrix. As in figure 6, below-left of the main diagonal shows the Monte Carlo values, and above-right shows the semi-analytic, inhomogeneous matrix result. The diagonal of the correlation matrix has been set to zero to enhance visibility of the off-diagonal elements. *Bottom right:* a direct comparison of the first off-diagonal bins ( $b' = b + 1$ ) in the average correlation matrix reveals  $\sim 5\%$ -level correlations in the simulations that are captured by the semi-analytic covariance.



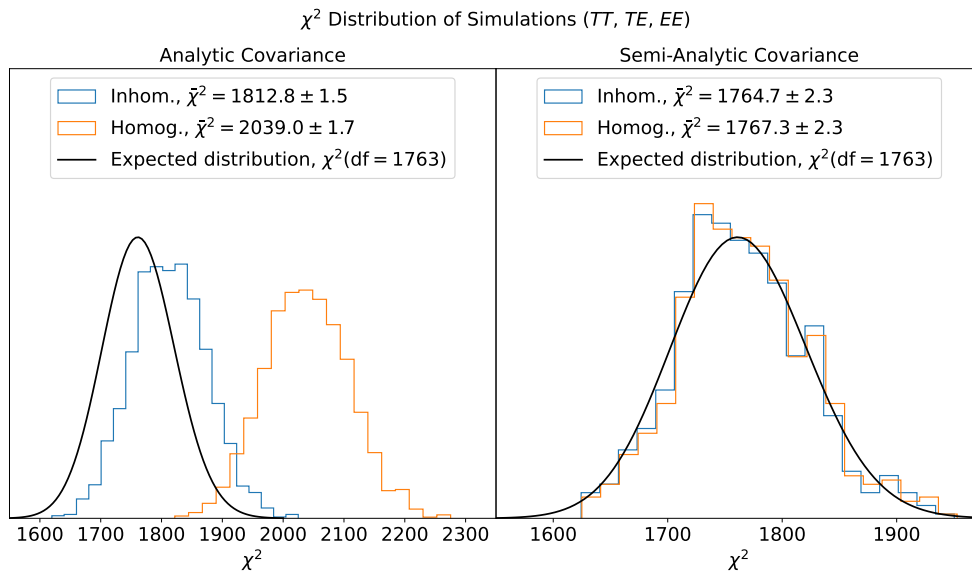
**Figure 8.** Ratios of the diagonal of the Monte Carlo covariance to the diagonal of the inhomogeneous analytic covariance, for a selection of array and polarization spectra. The PA5 f090 x PA6 f090  $TT$  spectra are the most signal-dominated; the PA6 f150 x PA6 f150  $EE$  spectra are the most noise-dominated. The grey shaded regions represent data not included in the likelihood from the scale cuts. The grey dashed line represents the amount of the observed excess Monte Carlo variance due to excess noise power in the simulations rather than a deficiency of the analytic covariance pipeline. It is an inhomogeneous analytic covariance matrix, but substitutes the data with a simulation when measuring the fiducial noise power spectra (see the discussion in section 5.2).

correlation matrix is defined as:

$$\rho_{ij} \equiv \Sigma_{ij} / \sqrt{\Sigma_{ii}\Sigma_{jj}}. \quad (5.1)$$

Figure 6 shows the similarity between the off-diagonal structures of the Monte Carlo and semi-analytic correlation matrices. We can easily discern the blocks associated with the 15  $TT$ , 16  $TE$ , and 10  $EE$  cross-array spectra in the data vector by their prominent bin-wise diagonals for each block.

We provide a more detailed view of the matrix structure in figure 7. The main diagonal of the  $EE$  blocks of  $\Sigma$  shows considerable dynamic range in simulations, spanning three orders of magnitude. Likewise, the bin-wise diagonals ( $b' = b$ ) of the first off-diagonal  $EE$  blocks ( $\beta' = \beta + 1$ ) of the correlation matrix also vary from  $\sim 80\%$  to near-zero correlation depending on the bin. In either case, the structures are well-modeled by the semi-analytic covariance. Finally, we average over all 100  $EE$  blocks in the covariance, elucidating that



**Figure 9.** Distributions of the  $d_{\text{sim}}^2$  (equation (5.3)) statistic for the simulated data vectors using different covariance matrices, after applying scale cuts and restricting to  $TT$ ,  $TE$ , and  $EE$  spectra. *Left:* using the inhomogeneous and the homogeneous uncorrected analytic covariance matrices, and evaluating all 1,600 simulations in our ensemble. *Right:* after the simulation-based correction (based on 1,000 simulations), and evaluating  $d_{\text{sim}}^2$  using the remaining 600 simulations. In both cases, the black line gives the expected distribution given the number of degrees-of-freedom (the length of the data vectors).

adjacent bins ( $b' = b + 1$ ) are also correlated, though only at the few-percent level. Otherwise, farther bin-wise off-diagonals ( $|b' - b| \geq 2$ ) are broadly consistent with noise for all blocks. We quantify that assessment as follows. Assuming zero population correlation, the variance of the Monte Carlo estimator for  $\rho_{ij}$  is:

$$\langle \rho_{ij}^2 \rangle \approx 1/n_{\text{sim}} \quad (5.2)$$

in the limit of a large number of simulations,  $n_{\text{sim}}$  [41].<sup>16</sup> For  $n_{\text{sim}} = 1,600$ , we would expect  $\bar{\rho}_{ij} = 0$  and  $\sigma(\rho_{ij}) = 2.5\%$ . We find  $\bar{\rho}_{ij} = -0.01\%$  and  $\sigma(\rho_{ij}) = 2.50\%$  (rounded to two decimals) for the  $|b' - b| \geq 2$  off-diagonals of the entire correlation matrix (figure 6), consistent with statistical scatter about zero correlation. Accordingly, the semi-analytic matrix predicts negligible correlation for these elements. In summary, the structures of the Monte Carlo and semi-analytic covariance matrices are consistent.

In addition to their off-diagonals, we compare a subset of the Monte Carlo and analytic matrix diagonals in figure 8. Here we are evaluating the uncorrected, inhomogeneous analytic covariance that results from section 4.1 alone. We do so to check the magnitude of any pipeline approximations breaking down, before the simulations correct for those approximations. As in figure 7, we again find percent-level agreement with the Monte Carlo matrix within the DR6 scale cuts. This finding holds across high and low signal-to-noise array-crosses (PA5 f090 x PA6 f090 vs. PA6 f150 x PA6 f150) and polarizations. In comparison to *Planck* [see appendix

<sup>16</sup>While the Monte Carlo covariance is Wishart distributed, the Monte Carlo correlation is more complicated. However, in this limit, they give the same result.

C.1.4 of [62], we do not find clear evidence for coherent  $\sim 10\%$ -level NKA violations; however, our bin resolution makes this determination difficult. The most significant deviations are seen at large scales in low signal-to-noise polarization blocks. As noted in section 4.3, however, these are due to excess power in the *simulations* entering the Monte Carlo matrix, not a breakdown of the analytic matrix pipeline. We discuss this aspect of the noise simulations further in section 5.2 and appendix C.1. Otherwise, we conclude that the Monte Carlo and analytic matrix diagonals are consistent at the percent-level even prior to the simulation-based correction. For example, of the diagonals shown in figure 8, the largest mean difference between the Monte Carlo and analytic covariance is 2.6% (PA6 f150 *EE*), of which 0.6% is due to excess power in the simulations.

Lastly, we test the uncorrected and corrected analytic covariance matrix by aggregating the  $\chi^2$  distribution of realistic simulations. For each simulation in the ensemble, we compute the squared Mahalanobis distance [55] of each simulation:

$$d_{\text{sim}}^2 \equiv (\hat{\mathbf{C}}_{\text{sim}} - \mathbf{C}_{th})^T \boldsymbol{\Sigma}^{-1} (\hat{\mathbf{C}}_{\text{sim}} - \mathbf{C}_{th}), \quad (5.3)$$

where  $\hat{\mathbf{C}}_{\text{sim}}$  are the reconstructed data vectors of the simulation,  $\mathbf{C}_{th}$  are the theoretical data vectors assuming the fiducial signal spectra of section 4.1.3, and  $\boldsymbol{\Sigma}$  is the trial covariance matrix. If the  $\hat{\mathbf{C}}_{\text{sim}}$  are normally distributed with mean  $\mathbf{C}_{th}$  and covariance  $\boldsymbol{\Sigma}$ , then  $d_{\text{sim}}^2$  are  $\chi^2$  distributed with degrees-of-freedom given by the size of the vector. This test is comprehensive as  $d_{\text{sim}}^2$  is sensitive to the entire covariance, and  $\hat{\mathbf{C}}_{\text{sim}}$  is free from analytic approximations. Results are shown in figure 9. For the uncorrected analytic covariance, we find our new inhomogeneous matrix leads to a simulation  $\chi^2$  distribution that is  $\sim 2.8\%$  greater than nominal. In other words, on average the uncorrected analytic covariance underestimates the error-bar of simulations by only  $\sim 1.4\%$ . This is a significant finding: despite the ways in which the DR6 data and processing break the assumptions of the MASTER covariance framework discussed in section 3, an approximate, MASTER-compatible prescription can limit bias to the percent-level. Consistent with the findings of ref. [7], the homogeneous matrix is  $\sim 15.7\%$  discrepant with simulations. Importantly, after applying the simulation-based correction of section 4.3 to each matrix, the resulting  $\chi^2$  distributions become consistent with the expected distribution at the  $\sim 0.3\%$ -level.<sup>17</sup> Thus, our simulation-based correction method is effective, even for  $\sim 16\%$ -level mismodeling of the input analytic covariance. We conclude that the semi-analytic covariance matrix for ACT DR6, whether using the inhomogeneous or homogeneous prescription, achieves satisfactory performance in the context of the challenging DR6 data properties.

## 5.2 Monte Carlo matrix validation

The results of section 5.1 assume the Monte Carlo covariance represents an unbiased estimate of the true covariance. Here, we investigate the validity of this assumption, focusing on the accuracy of the noise contribution. We briefly compared the recovered noise power spectra

<sup>17</sup>To obtain unbiased distributions, the covariance matrix and simulations in equation (5.3) must be statistically independent. We achieve this for the semi-analytic covariance matrices by using only 1,000 simulations to correct the covariance, and the remaining 600 to evaluate the  $d_{\text{sim}}^2$  statistics. We confirmed that the simulation-based correction is converged for the reduced number of simulations.

between simulations and the data in figures 5 and 8, noting that the large-scale, polarized noise power spectra in simulations tend to be larger than those of the data. The effect on the covariance is captured by the grey dashed line in figure 8, showing the uncorrected analytic covariance where, rather than measuring the fiducial noise power spectrum from the data, we replace the data with a noise simulation in section 4.1.4. This accounts for the mismatch between the Monte Carlo and analytic matrix that is due only to the simulation noise power spectra not matching that of the data, rather than due to a breakdown of the covariance pipeline approximations. We see that most of the low- $\ell$  Monte Carlo excess is due to this issue in the noise simulations, and is larger for covariance blocks that are more noise-dominated. This bias has a negligible effect given the DR6 scale cuts; at most, the error-bars of the  $\sim 3$  largest-scale  $EE$  bins above the scale cuts are  $\sim 1 - 2\%$  inflated. We discuss the origin of this issue in appendix C.1, and note it can be improved at larger scales for future analyses.

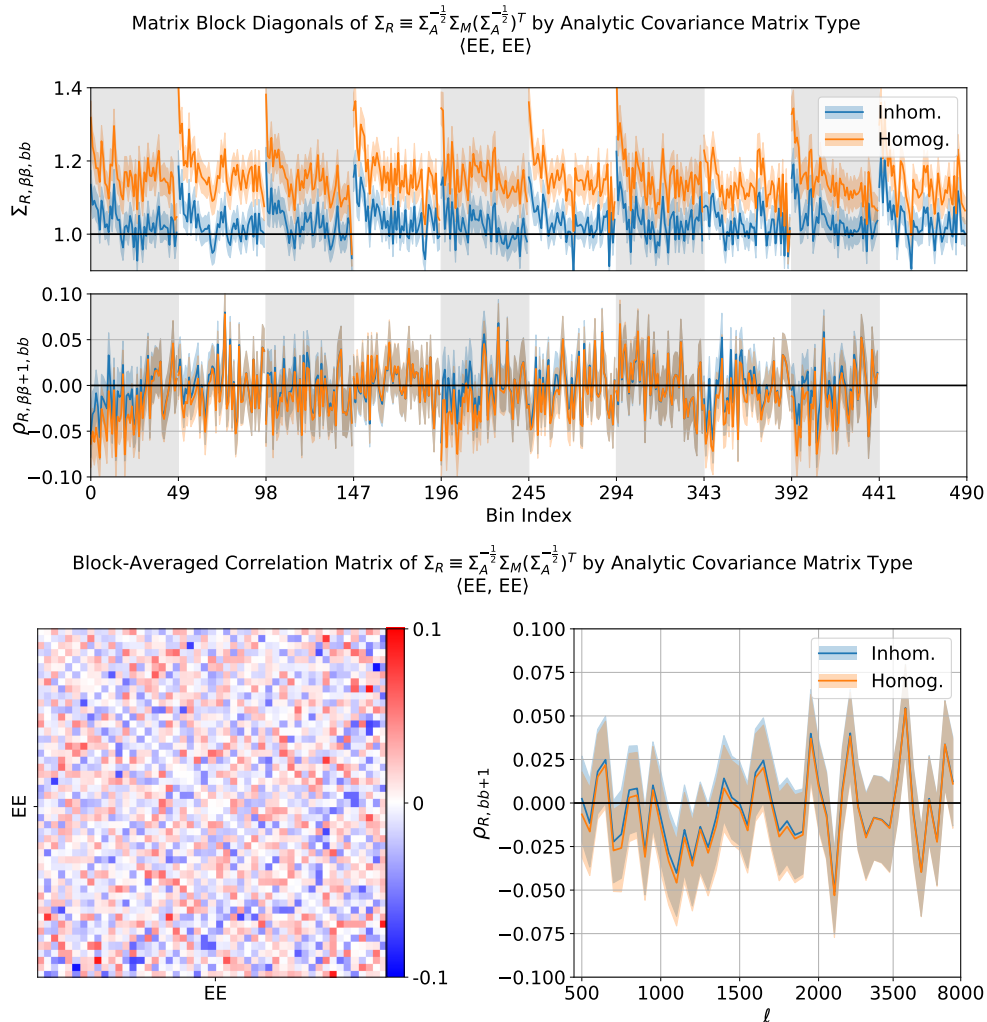
To attempt to quantify the importance of the spatially-varying anisotropy pattern, which cannot be easily accommodated in the MASTER covariance framework, we compare Monte Carlo matrices constructed from both ACT noise models. This test is well-motivated: the tiled model assumes the noise is diagonal in 2D Fourier space (such that the noise is close to diagonal in spherical harmonic space), whereas the directional wavelet noise model assumes long-range correlations in 2D Fourier space [for more detail, see 7]. We draw an ensemble of 600 directional wavelet-based simulations, and evaluate their  $\chi^2$  distribution against the tile-based semi-analytic covariance matrix, finding good agreement: the wavelet-based ensemble is only  $\sim 0.7\%$  different for the  $TT$ ,  $TE$ , and  $EE$  spectra. A stronger test selects only for the  $TB$ ,  $EB$ , and  $BB$  spectra, since these spectra are noise-dominated. Notably, we find a similar level of agreement:  $\sim 0.6\%$ .<sup>18</sup> Thus, we cannot detect any significant change in the Monte Carlo covariance, even when allowing for long-range noise correlations over angular scales.

### 5.3 Discussion of semi-analytic matrix performance

Here, we discuss why both the homogeneous and inhomogeneous prescriptions capture the covariance equally well after applying the simulation-based correction. The right panel of figure 9 suggests that the assumptions of our simulation-based correction method are reasonably met in both cases: that a *rotated* Monte Carlo covariance matrix,  $\Sigma_R$  (as defined in equation (4.7)), is approximately diagonal, and that its diagonal is approximately smooth. We test these predictions in figure 10. As expected from figure 9, the inhomogeneous analytic prescription underestimates the approximate Monte Carlo eigenspectrum by a few percent, as measured by the main diagonal of  $\Sigma_R$ , and in a smooth fashion easily handled by the simulation-based correction. The homogeneous analytic prescription underestimates the approximate Monte Carlo eigenspectrum by  $\sim 10 - 20\%$ , although with some prominent features extending as high as  $\sim 30\%$ . Evidently, these features are still easily smoothed by the Gaussian processes in section 4.3.

As discussed in section 4.3, figure 10 reveals  $\Sigma_R$  is not perfectly diagonal within statistical scatter: its off-diagonal blocks ( $\beta' \neq \beta$ ) contain percent-level correlations along their bin-wise

<sup>18</sup>In both these cases, we apply the lower bandlimit ( $\ell_{\max} = 5,400$ ) of the directional wavelet model to both the simulations and the covariance matrix. Doing so to a set of tiled simulations yields the same  $\sim 0.7\%$  (for  $TT$ ,  $TE$ , and  $EE$ ) and  $\sim 0.6\%$  (for  $TB$ ,  $EB$ , and  $BB$ ) difference, indicating that even these small discrepancies are due to the bandlimit, not the noise model.



**Figure 10.** *Top:* comparison of the rotated Monte Carlo matrices,  $\Sigma_R$  (equation (4.7)), when using the inhomogeneous (blue) or homogeneous (orange) analytic covariance matrices to perform the rotation. The covariance main diagonal elements are close to one, while the correlation bin-wise diagonals are small, but nonzero, for off-diagonal blocks. The matrix values can be directly compared to the first two rows of figure 7, showing the suppressed structure and dynamic range of  $\Sigma_R$  relative to  $\Sigma$ . *Bottom left:* the correlation matrix after averaging over all 100 blocks of the  $EE$  part of  $\Sigma_R$ . Below-left (above-right) of the main diagonal shows the result using the homogeneous (inhomogeneous) analytic matrix. The diagonal of the correlation matrix has been set to zero. *Bottom right:* the first off-diagonal bins ( $b' = b + 1$ ) in the average  $\Sigma_R$   $EE$  correlation matrix are consistent with each other and with zero for each analytic prescription, also unlike in figure 7.

diagonals ( $b' = b$ ). As expected, these off-diagonals are slightly more pronounced for the homogeneous matrix prescription, although still small. In either case, they are very smooth, and are thus easily fit by a Gaussian process (we explore the importance of accounting for these small off-diagonals in appendix D). Most importantly, for either analytic prescription,  $\Sigma_R$  contains significantly less structure and dynamic range than the raw covariance or correlation matrices in figure 7. Moreover, the bin-wise off-diagonal ( $b' = b + 1$ ) correlations, which are present at the  $\sim 5\%$ -level in the raw correlation matrix in figure 7, are absent for  $\Sigma_R$ , again regardless of the analytic prescription. Therefore, while the core assumptions motivating  $\Sigma_R$  are not perfect, they are good enough to enable our simple simulation-based correction scheme.

Nevertheless, a covariance structure dominated by block-diagonals does not guarantee satisfactory performance when applying any simulation-based matrix correction. For example, a conceptually similar simulation-based correction method to that presented in section 4.3 is to assume that the Monte Carlo and analytic matrices share correlation matrices ( $\rho_{ij}$ ) rather than eigenbases. The correction then takes the form of reweighting the analytic matrix diagonal using the Monte Carlo matrix while preserving its correlation matrix. Such a scheme was used in WMAP [72] and was suggested by ref. [35]. We also tested this method, where we performed the reweighting of the matrix diagonal using the same Gaussian process smoothing as we developed in section 4.3. We found this simulation-based correction method did not perform as well. Using our inhomogeneous analytic covariance, the corrected matrix resulted in a  $\chi^2$  distribution of simulations (as in the right panel of figure 9) that remained  $\sim 1.7\%$  too high, while using the homogeneous analytic covariance resulted in a distribution that remained  $\sim 4.9\%$  too high. Therefore, we conclude that the Monte Carlo and analytic covariance do not share exact correlation structures; rather, the assumption of a shared eigenbasis is more robust.

While not critical to our covariance validation, for the interested reader we also investigate the effect of point-source holes in the analysis mask (see figure 1) on the covariance matrix in appendix E.

## 6 Conclusion

We have developed a new power spectrum covariance matrix pipeline for the ACT DR6 data in response to the challenge the data pose to the MASTER covariance framework. While the resulting inhomogeneous prescription is similar to that presented in *Planck* [61, 62], unlike *Planck* the ACT data do not conform to the data model required in MASTER. Through our use of different effective spatial weights for signal and noise, we also depart from existing public covariance implementations. We find that our pipeline yields an analytic covariance matrix for the ACT DR6 data that is accurate to within  $\sim 3\%$  of a Monte Carlo covariance that incorporates realistic properties of the ACT data. While we observe  $\sim 16\%$ -level differences between an analogous homogeneous matrix and the Monte Carlo matrix, both analytic matrices achieve sub-percent agreement with the Monte Carlo covariance after applying a new simulation-based correction method. In comparison, a common correction method assuming that the analytic correlation matrix is accurate, while only correcting the covariance diagonal, results in biases as large as  $\sim 5\%$ . Our central result is that a semi-analytic covariance using

our simulation-based correction and either the inhomogeneous or homogeneous prescriptions is well-suited for use in the ACT DR6 likelihood.

This work has clarified the mechanisms behind the  $\sim 16\%$ -level discrepancies between the Monte Carlo and the homogeneous analytic covariance matrix. Accepting simulation-based corrections of that magnitude, without understanding what is driving them, may have limited our confidence in the final covariance matrix. Instead, we understand most of the salient features of the simulations at the covariance level: by making a set of *a priori* well-motivated changes to the pipeline, summarized in section 4.1.6, we achieve a discrepancy between the analytic and Monte Carlo covariance of less than 3%. This gives us sufficient confidence in the simulations to condition the homogeneous analytic covariance, even for corrections as large as  $\sim 16\%$ .

This work has implications for future large-aperture CMB experiments, such as the Simons Observatory (SO). A change to the map-based noise properties, or a change in power spectrum bin size, could result in a noticeable difference in performance between the homogeneous and inhomogeneous prescriptions, even after attempts at matrix conditioning. The homogeneous prescription is also more reliant on a well-converged Monte Carlo covariance, so this new prescription may allow for a smaller simulation ensemble. Between this new analytic prescription and continued simulation development, we anticipate this power spectrum covariance matrix will continue to meet requirements for next-generation CMB science programs.

## Acknowledgments

ZA and JD acknowledge support from NSF grant AST-2108126. This work was supported by a grant from the Simons Foundation (CCA 918271, PBL). The Flatiron Institute is supported by the Simons Foundation. CS acknowledges support from the Agencia Nacional de Investigación y Desarrollo (ANID) through Basal project FB210003. KM acknowledges support from the National Research Foundation of South Africa. EC and HJ acknowledge support from the Horizon 2020 ERC Starting Grant (Grant agreement No 849169); SG and EC also acknowledge support from STFC and UKRI (grant numbers ST/W002892/1 and ST/X006360/1). 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 operated 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. Computing for ACT was performed using the Princeton Research Computing resources at Princeton University, the National Energy Research Scientific Computing Center (NERSC), and 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. We thank the Republic of Chile for hosting ACT in the

northern Atacama, and the local indigenous Licanantay communities whom we follow in observing and learning from the night sky.

Some of the results in this paper have been derived using the `healpy` and `HEALPix` package [33, 74]. Other software used in this paper include the following: `astropy` [4–6], `cython` [10], `ducc` (<https://gitlab.mpcdf.mpg.de/mtr/ducc>), `h5py` [17], `libsharp` [66], `matplotlib` [42], `mnms` [7], `numba` [46], `numpy` [36], `pyfftw` [32], `pyyaml` (<https://pyyaml.org/>), and `scipy` [73].

## A More details on data model

We presented the ACT data model in section 3.1; for brevity, we omitted some pieces that we document here. These pieces do not fundamentally change the data description, but do complicate our map-level simulation somewhat beyond what is prescribed in equation (3.4). The simulations in section 4.2 follow the additional steps of this section.

The beam-convolved sky signal defined in equation (3.2) does not exactly correspond to the signal in the ACT maps. Instead, optical and readout inefficiencies result in an overall calibration error in the maps, and the map digitization — or “mapmaking” — introduces a pixel window function:

$$\mathbf{s} = \mathbf{A}\mathbf{F}^\dagger\mathbf{X}_p\mathbf{F}\mathbf{Y}\mathbf{B}\mathbf{S}^{\frac{1}{2}}\boldsymbol{\eta}_s^h, \quad (\text{A.1})$$

where  $\mathbf{F}^\dagger\mathbf{X}_p\mathbf{F}$  applies a pixel window function matrix  $\mathbf{X}_p$ , which is diagonal Fourier space, to the maps, and  $\mathbf{A}$  is a diagonal matrix containing the observational efficiency factor for each array and polarization. We account for the map noise “as is,” in that equation (3.3) models the noise in the map products themselves, and so it is arbitrary whether we include the calibration or pixel window in the noise model as long as we are consistent. We opt to exclude them.

As discussed in section 3.1, after the ACT maps are made, they are further processed as part of the power spectrum estimation pipeline (the ACT DR6 power spectrum paper will provide a more detailed explanation). In section 3.1, we omitted two steps in that processing. Firstly, measurement of the ACT observational efficiency in both temperature and polarization is made through spectrum-level comparisons against *Planck*. These corrections are defined and applied at the map level. In terms of our data model from section 3.1, this looks like constructing  $\mathbf{A}^{-1}$  and applying it to  $\mathbf{m}$ :  $\mathbf{A}^{-1}\mathbf{m}$ . Next, the pixel window function is divided out in Fourier space. In fact, because both the pixel window function and the Fourier filter are diagonal in Fourier space, both steps are performed simultaneously. Importantly, since these operations are slightly non-local in map space, prior to performing them the maps are masked with a broad, apodized sky mask which eliminates bright galactic regions and noisy pixels near the edge of the ACT footprint. Altogether, this applies the following matrix to our calibrated map model:  $\mathbf{F}^\dagger\mathbf{X}_f\mathbf{X}_p^{-1}\mathbf{F}\mathbf{W}_k$ . Here,  $\mathbf{W}_k$  is the sky mask applied before the Fourier operations,  $\mathbf{X}_p$  is the pixel window function, and  $\mathbf{X}_f$  is the pickup filter in Fourier space. Together with applying the analysis mask ( $\mathbf{W}$ ) as discussed in section 3.1, we have the following data model

of the processed ACT maps (again, in the case of the “tiled” noise model):

$$\begin{aligned}
 \mathbf{m}_i &\rightarrow \mathbf{W}\mathbf{F}^\dagger\mathbf{X}_f\mathbf{X}_p^{-1}\mathbf{F}\mathbf{W}_k\mathbf{A}^{-1}\mathbf{m}_i \\
 &= \mathbf{W}\mathbf{F}^\dagger\mathbf{X}_f\mathbf{X}_p^{-1}\mathbf{F}\mathbf{W}_k\mathbf{F}^\dagger\mathbf{X}_p\mathbf{F}\mathbf{Y}\mathbf{B}\mathbf{S}^{\frac{1}{2}}\boldsymbol{\eta}_s^h \\
 &\quad + \mathbf{W}\mathbf{F}^\dagger\mathbf{X}_f\mathbf{X}_p^{-1}\mathbf{F}\mathbf{W}_k\mathbf{A}^{-1}\boldsymbol{\sigma}_i\mathbf{Y}\mathbf{N}_i^{\frac{1}{2}}\mathbf{Y}^\dagger\boldsymbol{\Omega}\mathcal{T}^\dagger\mathbf{F}^\dagger\mathcal{N}_i^{\frac{1}{2}}\boldsymbol{\eta}_{n,i}^{\mathcal{T}f}.
 \end{aligned} \tag{A.2}$$

This is the full forward model for the DR6 simulated maps discussed in section 4.2.

In practice, equation (A.2) is too cumbersome to work with for the analytic covariance matrix, so we make further simplifying assumptions. We discussed one in section 3.3: the approximate noise model (equation (3.10)). Furthermore, since the isotropic noise power spectrum is measured from the data (see section 4.1.4), we can absorb the map calibration operation,  $\mathbf{A}^{-1}$ , into it. Then, for both the signal and noise terms, we assume that after applying the analysis mask,  $\mathbf{W}$ , edge effects from the Fourier operations on the  $\mathbf{W}_k$ -masked data are indeed negligible, such that we can eliminate  $\mathbf{W}_k$  from equation (3.4). Finally, since the pixel window function (and its inverse),  $\mathbf{X}_p$ , is almost isotropic, and the noise is already highly anisotropic, we also absorb  $\mathbf{X}_p^{-1}$  into the isotropic noise power spectrum of the approximate noise model. In other words, neglecting the anisotropy of  $\mathbf{X}_p^{-1}$  is a negligible addition to the data model error already incurred by the approximate noise model. Cancelling terms leads to the following simplified data model of the processed maps:

$$\mathbf{m}_i = \mathbf{W}(\mathbf{F}^\dagger\mathbf{X}_f\mathbf{F})\mathbf{Y}\mathbf{B}\mathbf{S}^{\frac{1}{2}}\boldsymbol{\eta}_s^h + \mathbf{W}(\mathbf{F}^\dagger\mathbf{X}_f\mathbf{F})\boldsymbol{\sigma}_i\boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{Y}\mathbf{N}_i^{\frac{1}{2}}\boldsymbol{\eta}_{n,i}^{\mathcal{T}f}, \tag{A.3}$$

where  $\mathbf{N}_i$  is measured *after* the map processing, and we have placed in parentheses the Fourier-space pickup filter (the  $\mathbf{F}^\dagger\mathbf{X}_f\mathbf{F}$  operator). Replacing the Fourier-space pickup filter with an effective transfer function in harmonic space, as described in section 3.3, yields equation (3.11).

We also show the origin of the  $\boldsymbol{\Omega}^{\frac{1}{2}}$  factor in the effective noise weight when going from equation (3.9) to equation (3.10). The factor appears in the literature as early as ref. [24], and recurs throughout the *Planck* covariance matrices, but its whereabouts have not been explicitly discussed. A field that follows the model of equation (3.9) has zero mean and is Gaussian, so its statistics are fully specified by its covariance. This is given by:

$$\begin{aligned}
 \langle \mathbf{n}_i\mathbf{n}_i^\dagger \rangle &= \boldsymbol{\sigma}_i\mathbf{Y}\mathbf{N}_i^{\frac{1}{2}}\mathbf{Y}^\dagger\boldsymbol{\Omega}\langle \boldsymbol{\eta}_{n,i}^m\boldsymbol{\eta}_{n,i}^{m,\dagger} \rangle\boldsymbol{\Omega}\mathbf{Y}(\mathbf{N}_i^{\frac{1}{2}})^T\mathbf{Y}^\dagger\boldsymbol{\sigma}_i \\
 &= \boldsymbol{\sigma}_i\mathbf{Y}\mathbf{N}_i^{\frac{1}{2}}\mathbf{Y}^\dagger\boldsymbol{\Omega}^2\mathbf{Y}(\mathbf{N}_i^{\frac{1}{2}})^T\mathbf{Y}^\dagger\boldsymbol{\sigma}_i \\
 &\approx \boldsymbol{\sigma}_i\boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{Y}\mathbf{N}_i^{\frac{1}{2}}(\mathbf{Y}^\dagger\boldsymbol{\Omega}\mathbf{Y})(\mathbf{N}_i^{\frac{1}{2}})^T\mathbf{Y}^\dagger\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{\sigma}_i \\
 &= \boldsymbol{\sigma}_i\boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{Y}\mathbf{N}_i\mathbf{Y}^\dagger\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{\sigma}_i,
 \end{aligned} \tag{A.4}$$

where the following definitions and properties were used: that  $\boldsymbol{\eta}_{n,i}^m$  is a white-noise vector in map space, that  $\boldsymbol{\Omega}^{\frac{1}{2}}$  is a smooth map that nearly commutes with the harmonic operations in terms like  $\mathbf{Y}\mathbf{N}_i^{\frac{1}{2}}\mathbf{Y}^\dagger$ , and that  $\mathbf{Y}^\dagger\boldsymbol{\Omega}\mathbf{Y}$  is the identity. The covariance in the last line of equation (A.4) is the same as that for a field following the model of equation (3.10). Therefore, fields following equations (3.9) and (3.10) are nearly identically distributed.

## B Analytical pseudospectrum covariance matrices

In this section, we derive equation (4.2) and give our expressions for the coupling ‘‘spin,’’  $\beta$ . We also motivate our use of an isotropic transfer function to approximately model the Fourier-space filter, as described in section 4.1.2, and the steps to measure each of its components.

### B.1 Covariance matrices

The structure of the pseudospectrum covariance matrix term in equation (4.2) is similar to that of ref. [62] and ref. [59], but here we provide an explicit derivation for reference. We work in the temperature-only case, but the derivation is analogous for polarization [see e.g., 11, 12, 14, 18, for the appropriate substitutions].

We start by defining a few mathematical tools. First, we can write equation (3.5) in harmonic space as [see e.g., 39]:

$$\tilde{a}_{\ell m} = \sum_{\ell' m'} K_{\ell m, \ell' m'}(w) a_{\ell' m'}, \quad (\text{B.1})$$

where

$$K_{\ell m, \ell' m'}(w) \equiv \int dx Y_{\ell' m'}(x) w(x) Y_{\ell m}^*(x). \quad (\text{B.2})$$

The  $K$  matrices have the following property [see e.g., 18]:

$$K_{\ell_1 m_1, \ell_2 m_2}(yz) = \sum_{\ell_3 m_3} K_{\ell_1 m_1, \ell_3 m_3}(y) K_{\ell_3 m_3, \ell_2 m_2}(z) = \sum_{\ell_3 m_3} K_{\ell_1 m_1, \ell_3 m_3}(y) K_{\ell_2 m_2, \ell_3 m_3}^*(z), \quad (\text{B.3})$$

where  $yz$  is the product of the two masks  $y$  and  $z$  in map space. Finally, we need the definition of the coupling matrices,  $\Xi_{\ell \ell'}(y, z)$ , from equation (3.6):

$$\Xi_{\ell \ell'}(y, z) \equiv \frac{1}{2\ell + 1} \frac{1}{2\ell' + 1} \sum_{mm'} K_{\ell m, \ell' m'}(y) K_{\ell m, \ell' m'}^*(z). \quad (\text{B.4})$$

As noted in ref. [39], equation (B.4) can be calculated in terms of the cross-power spectrum of the masks  $y$  and  $z$  and the Wigner 3-j symbols in  $\mathcal{O}(\ell_{\max}^3)$ -time.

Next, consider four instances of an isotropic field,  $a$ , and mask,  $w$ , that follow the model of equation (3.8). We label each instance  $i$ ,  $j$ ,  $p$ , and  $q$ . The fields are fully specified by their cross-power spectra,  $C_\ell^{yz}$ , where  $yz$  is a pair of labels:

$$\langle a_{\ell m}^y a_{\ell' m'}^z \rangle_s = C_\ell^{yz} \delta_{\ell \ell'} \delta_{mm'}. \quad (\text{B.5})$$

The pseudospectrum estimator for a pair of these fields (equation (3.6)) is  $\hat{C}_\ell^{yz}$ , and the covariance of  $i$ ,  $j$ ,  $p$ , and  $q$  is:

$$\begin{aligned} \tilde{\Sigma}_{\ell \ell'}^{i, j, p, q} &\equiv \langle \hat{C}_\ell^{ij} \hat{C}_{\ell'}^{pq} \rangle - \langle \hat{C}_\ell^{ij} \rangle \langle \hat{C}_{\ell'}^{pq} \rangle \\ &= \frac{1}{2\ell + 1} \frac{1}{2\ell' + 1} \sum_{m, m'} \langle \tilde{a}_{\ell m}^i \tilde{a}_{\ell m}^{j*} \tilde{a}_{\ell' m'}^p \tilde{a}_{\ell' m'}^{q*} \rangle - \langle \tilde{a}_{\ell m}^i \tilde{a}_{\ell m}^{j*} \rangle \langle \tilde{a}_{\ell' m'}^p \tilde{a}_{\ell' m'}^{q*} \rangle. \end{aligned} \quad (\text{B.6})$$

Because  $\tilde{a}$  is Gaussian, we expand the product of four fields using Wick's theorem. Also using the reality of the fields, equation (B.6) becomes:

$$\tilde{\Sigma}_{\ell\ell'}^{i,j,p,q} = \frac{1}{2\ell+1} \frac{1}{2\ell'+1} \sum_{m,m'} \langle \tilde{a}_{\ell m}^i \tilde{a}_{\ell' m'}^{p*} \rangle \langle \tilde{a}_{\ell m}^{j*} \tilde{a}_{\ell' m'}^q \rangle + \langle \tilde{a}_{\ell m}^i \tilde{a}_{\ell' m'}^{q*} \rangle \langle \tilde{a}_{\ell m}^{j*} \tilde{a}_{\ell' m'}^p \rangle. \quad (\text{B.7})$$

Combining with equations (B.1) and (B.5), we get:

$$\begin{aligned} \tilde{\Sigma}_{\ell\ell'}^{i,j,p,q} &= \frac{1}{2\ell+1} \frac{1}{2\ell'+1} \sum_{m,m'} \sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w^i) K_{\ell' m', \ell_1 m_1}^*(w^p) C_{\ell_1}^{ip} \sum_{\ell_3 m_3} K_{\ell m, \ell_3 m_3}^*(w^j) K_{\ell' m', \ell_3 m_3}(w^q) C_{\ell_3}^{jq} \\ &+ \sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w^i) K_{\ell' m', \ell_1 m_1}^*(w^q) C_{\ell_1}^{iq} \sum_{\ell_3 m_3} K_{\ell m, \ell_3 m_3}^*(w^j) K_{\ell' m', \ell_3 m_3}(w^p) C_{\ell_3}^{jp}. \end{aligned} \quad (\text{B.8})$$

Each term like  $\sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w^i) K_{\ell' m', \ell_1 m_1}^*(w^p) C_{\ell_1}^{ip}$  is difficult to compute directly, but if we assume the NKA, then we can move  $C_{\ell_1}^{ip}$  out of the sum over  $\ell_1$ , yielding:

$$\sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w^i) K_{\ell' m', \ell_1 m_1}^*(w^p) C_{\ell_1}^{ip} \rightarrow C_{(\ell, \ell')}^{ip} \sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w^i) K_{\ell' m', \ell_1 m_1}^*(w^p), \quad (\text{B.9})$$

where  $(\ell, \ell')$  denotes some symmetric function of  $\ell$  and  $\ell'$  [24]. Finally, combining equations (B.3), (B.4), (B.8), and (B.9) yields:

$$\tilde{\Sigma}_{\ell\ell'}^{i,j,p,q} = C_{(\ell, \ell')}^{ip} C_{(\ell, \ell')}^{jq} \Xi_{\ell\ell'}(w^i w^p, w^j w^q) + C_{(\ell, \ell')}^{iq} C_{(\ell, \ell')}^{jp} \Xi_{\ell\ell'}(w^i w^q, w^j w^p). \quad (\text{B.10})$$

In the case of equation (4.1), we have additive signal (fields  $s$ , masks  $u$ , and power spectra  $S$ ) and noise (fields  $n$ , masks  $v$ , and power spectra  $N$ ) that are uncorrelated. Substituting equation (4.1) for  $\tilde{a}$  in equation (B.7), and expanding into signal and noise, would yield 16  $(ip, jq)$  signal-noise cross terms and 16  $(iq, jp)$  signal-noise cross terms, but assuming the signal and noise are uncorrelated reduces this to four of each term, in particular:

$$\begin{aligned} \tilde{\Sigma}_{\ell\ell'}^{i,j,p,q} &= S_{(\ell, \ell')}^{ip} S_{(\ell, \ell')}^{jq} \Xi_{\ell\ell'}(u^i u^p, u^j u^q) + S_{(\ell, \ell')}^{ip} N_{(\ell, \ell')}^{jq} \Xi_{\ell\ell'}(u^i u^p, v^j v^q) \\ &+ N_{(\ell, \ell')}^{ip} S_{(\ell, \ell')}^{jq} \Xi_{\ell\ell'}(v^i v^p, u^j u^q) + N_{(\ell, \ell')}^{ip} N_{(\ell, \ell')}^{jq} \Xi_{\ell\ell'}(v^i v^p, v^j v^q) \\ &+ S_{(\ell, \ell')}^{iq} S_{(\ell, \ell')}^{jp} \Xi_{\ell\ell'}(u^i u^q, u^j u^p) + S_{(\ell, \ell')}^{iq} N_{(\ell, \ell')}^{jp} \Xi_{\ell\ell'}(u^i u^q, v^j v^p) \\ &+ N_{(\ell, \ell')}^{iq} S_{(\ell, \ell')}^{jp} \Xi_{\ell\ell'}(v^i v^q, u^j u^p) + N_{(\ell, \ell')}^{iq} N_{(\ell, \ell')}^{jp} \Xi_{\ell\ell'}(v^i v^q, v^j v^p), \end{aligned} \quad (\text{B.11})$$

which is equivalent to equation (4.2).

To fully specify our covariance matrix prescription in equation (4.2), we provide the  $\beta$  mapping from field polarizations to coupling spins in table 2. This follows from the INKA, as well as the assumption that mask gradients can be neglected [18], such that “--” couplings are small compared to “++” couplings.

## B.2 More details on Fourier filter

In this section, we first examine the effect of a Fourier-space filter analytically by considering it in harmonic space, and then discuss how we construct our approximation for the Fourier-space filter described in section 4.1.2.

$\beta(AB, CD)$	$AB, CD$
00	TT, TT
	TT, TP + three permutations
	TP, TP + three permutations
0+	TT, PP + one permutation
	TP, PP + 3 permutations
++	PP, PP
--	Not used

**Table 2.** Coupling spins as a function of input field polarizations.  $P$  refers to either  $E$  or  $B$ . Note, there are 16 possible polarization permutations, and in no case do we use the  $--$  coupling.

A filter applied to fields in Fourier (or harmonic) space has analogous effects on the power spectrum and covariance matrix as a mask applied in map space. For example, the different manifestation of the mask at the spectrum (two-point) level and covariance (four-point) level results in the factor of the mask sky fraction,  $f_{sky}$ , in the denominator of the leading-order power spectrum covariance matrix [43]. To derive a similar effect for a filter, first consider a filter applied to fields in harmonic space in the absence of any mask. The field has the following data model:

$$\tilde{a}_{\ell m} = f_{\ell m} C_{\ell}^{\frac{1}{2}} \eta_{\ell m}, \quad (\text{B.12})$$

where the filter  $f_{\ell m}$  can be anisotropic ( $m$ -dependent). In expectation, the ‘‘pseudospectrum’’ of  $\tilde{a}$  is then related to the power spectrum  $C_{\ell}$  as:

$$\langle \hat{C}_{\ell} \rangle = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \langle \tilde{a}_{\ell m} \tilde{a}_{\ell m}^* \rangle = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} f_{\ell m}^2 C_{\ell} \equiv t_{\ell, 2pt} C_{\ell}, \quad (\text{B.13})$$

where we have defined the two-point isotropic transfer function:

$$t_{\ell, 2pt} \equiv \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} f_{\ell m}^2. \quad (\text{B.14})$$

The covariance of the pseudospectrum is given by equation (B.7):

$$\begin{aligned} \tilde{\Sigma}_{\ell\ell'} &\equiv \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{m, m'} \langle |\tilde{a}_{\ell m} \tilde{a}_{\ell' m'}^*|^2 \rangle = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{m, m'} |f_{\ell m} f_{\ell' m'} C_{\ell}^{\frac{1}{2}} C_{\ell'}^{\frac{1}{2}} \langle \delta_{\ell\ell'} \delta_{mm'} \rangle|^2 \\ &= \frac{2\delta_{\ell\ell'}}{(2\ell + 1)^2} \sum_m f_{\ell m}^4 C_{\ell}^2 \\ &\equiv \frac{2t_{\ell, 4pt} C_{\ell}^2 \delta_{\ell\ell'}}{2\ell + 1}, \end{aligned} \quad (\text{B.15})$$

where we have defined the four-point isotropic transfer function:

$$t_{\ell, 4pt} \equiv \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} f_{\ell m}^4. \quad (\text{B.16})$$

Then the covariance matrix of the *power spectrum* is given by:

$$\Sigma_{\ell\ell'} = \frac{1}{t_{\ell,2pt}t_{\ell',2pt}}\tilde{\Sigma}_{\ell\ell'} = \frac{2t_{\ell,4pt}C_{\ell}^2\delta_{\ell\ell'}}{t_{\ell,2pt}^2(2\ell+1)}. \quad (\text{B.17})$$

The filter factor  $t_{\ell,4pt}/t_{\ell,2pt}^2$  has non-trivial behavior depending on the character of  $f_{\ell m}$ . If  $f_{\ell m}$  is actually isotropic — with no  $m$ -dependence, like a beam — then  $t_{\ell,4pt}/t_{\ell,2pt}^2 = 1$ . If  $f_{\ell m}$  is binary — equal only to 0 or 1 — then  $t_{\ell,4pt} = t_{\ell,2pt} \equiv t_{\ell}$  and  $t_{\ell,4pt}/t_{\ell,2pt}^2 = 1/t_{\ell}$ , where  $t_{\ell} \equiv 1/(2\ell+1)\sum_m f_{\ell m}$  is the average value of  $f_{\ell m}$  over  $m$ . If  $f_{\ell m}$  has any anisotropy, then it is less than 1. In this case, we see an analogy to a mask: the covariance matrix increases as  $1/t_{\ell}$ .

When there is a mask in addition to a harmonic-space filter, these expressions become inexact. The pseudospectrum is then related to the power spectrum as:

$$\langle \hat{C}_{\ell} \rangle = \frac{1}{2\ell+1} \sum_{\ell'} C_{\ell'} \sum_{mm'} K_{\ell m, \ell' m'}(w) K_{\ell m, \ell' m'}^*(w) f_{\ell' m'}^2. \quad (\text{B.18})$$

Ordinarily, at this point, the MASTER formalism would use equation (B.4), but if  $f_{\ell m}$  is anisotropic, this simplification is no longer possible. To recover the MASTER formalism, we approximate  $f_{\ell m}^2$  as being equal to its isotropic average, raised to some exponent,  $\alpha_{2pt}$ :

$$\sum_{mm'} K_{\ell m, \ell' m'}(w) K_{\ell m, \ell' m'}^*(w) f_{\ell' m'}^2 \rightarrow t_{\ell}^{\alpha_{2pt}} \sum_{mm'} K_{\ell m, \ell' m'}(w) K_{\ell m, \ell' m'}^*(w), \quad (\text{B.19})$$

where:

$$t_{\ell} \equiv \frac{1}{2\ell+1} \sum_m f_{\ell m}^2. \quad (\text{B.20})$$

To be sure, the substitution in equation (B.19) is an ansatz meant to approximately capture the actual expression. The ansatz assumes the effect of the harmonic-space filter is entirely imparted on the underlying power spectrum of the filtered field, with no effect on the mode-coupling matrix  $M_{\ell\ell'}$ . Overall, we find this ansatz works quite well, and is independent of whether equation (B.20) is binned, indicating that there are not significant changes to the mode-coupling structure. We then have the following for the pseudospectrum:

$$\langle \hat{C}_{\ell} \rangle = \sum_{\ell'} M_{\ell\ell'} t_{\ell}^{\alpha_{2pt}} C_{\ell'}, \quad (\text{B.21})$$

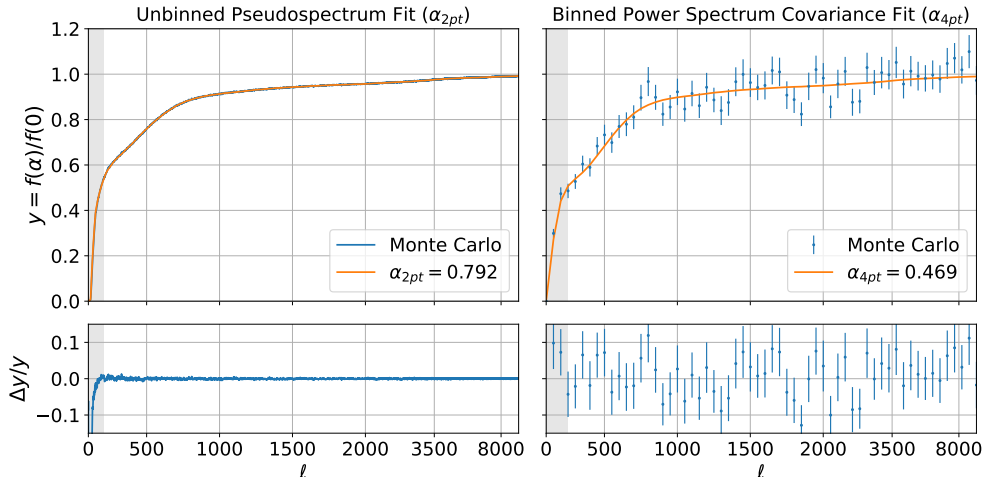
as discussed in section 4.1.2.

For the pseudospectrum covariance, we go back to equation (B.9), which in the presence of a harmonic-space filter becomes:

$$\sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w) K_{\ell' m', \ell_1 m_1}^*(w) f_{\ell_1 m_1}^2 C_{\ell_1} \rightarrow (t^{\alpha_{4pt}} C)_{(\ell, \ell')} \sum_{\ell_1 m_1} K_{\ell m, \ell_1 m_1}(w) K_{\ell' m', \ell_1 m_1}^*(w), \quad (\text{B.22})$$

where we have made a similar ansatz as equation (B.19), but as it involves a different sum (over  $\ell$  and  $m$ , not just over  $m$ ), we assign a different exponent to  $t_{\ell}$  —  $\alpha_{4pt}$  — and we still use the NKA:  $(t^{\alpha_{4pt}} C)_{(\ell, \ell')}$  denotes a symmetric function of  $t_{\ell}^{\alpha_{4pt}} C_{\ell}$  and  $t_{\ell'}^{\alpha_{4pt}} C_{\ell'}$ . The pseudospectrum covariance matrix (equation (B.10)) then becomes:

$$\tilde{\Sigma}_{\ell\ell'} = 2(t^{\alpha_{4pt}} C)_{(\ell, \ell')}^2 \Xi_{\ell\ell'}(w^2, w^2). \quad (\text{B.23})$$



**Figure 11.** *Left:* fitting the model of equation (B.21) to simulated pseudospectra by optimizing  $\alpha_{2pt}$ . *Right:* fitting the model of equation (B.24) to simulated Monte Carlo covariance diagonals by optimizing  $\alpha_{4pt}$ . In both cases, the fits are performed (and plotted) after normalizing the model by the case of  $\alpha = 0$ . The model is denoted by  $f$ , and the normalized model is denoted by  $y$ . Bottom panels give the residuals relative to the fitted model. The fit only uses scales where  $t_\ell > 0.5$ ; the excluded large-scales are shaded in grey.

We have made the same assumption that the effect of the harmonic-space filter is only on the underlying power spectrum and not on the coupling matrix. We find that this ansatz performs reasonably well, but unlike equation (B.21), we find the binning does matter, and also find the preferred value of  $\alpha_{4pt}$  changes when using an expression for the binned *power spectrum* covariance matrix:

$$\Sigma_{bb'} = \sum_{\ell} Q_{b\ell} \sum_{\ell'} Q_{b'\ell'} \tilde{\Sigma}_{\ell\ell'} = 2 \sum_{\ell} Q_{b\ell} \sum_{\ell'} Q_{b'\ell'} (t^{\alpha_{4pt}} C)_{(\ell,\ell')}^2 \Xi_{\ell\ell'}(w^2, w^2), \quad (\text{B.24})$$

where, unlike equation (4.4), here we use a simplified, diagonal  $\mathbf{U}$  matrix with diagonal equal to  $1/t_\ell^{\alpha_{2pt}}$ , following equation (B.21). These findings indicate that this ansatz is not optimal, but given our simulation-based correction, it is sufficient.

In reality, our filter lives in Fourier space, not harmonic space, but we follow the setup of the preceding discussion regardless in building a MASTER-compatible filter approximation. Since our analysis includes a mask, we fit for each quantity of interest —  $t_\ell$ ,  $\alpha_{2pt}$ , and  $\alpha_{4pt}$  — using simulations. We first draw 50 full-sky, white-noise simulations, filter them in Fourier space, and measure the transfer function template,  $t_\ell$ , directly via equation (B.13). We have a prior assumption that  $t_\ell$  should be smooth, so we can low-pass filter the noisy Monte Carlo estimate of  $t_\ell$  to get a better measurement. We use a Savitzky-Golay filter (implemented in `scipy.signal.savgol_filter`) with a window length of 100 and a polynomial order of 4. As the basis of our fit for  $\alpha_{2pt}$  and  $\alpha_{4pt}$ , we next draw an ensemble of 500 simulations that include masking. These simulations are drawn according to the following data model for a single scalar field:

$$\mathbf{a} = \mathbf{W}(\mathbf{F}^\dagger \mathbf{X}_f \mathbf{F}) \mathbf{W}_k \mathbf{YBC}^{\frac{1}{2}} \boldsymbol{\eta}^h, \quad (\text{B.25})$$

which, compared to equation (A.2), omits all anisotropic features other than the Fourier filter (i.e., intrinsic noise anisotropies and the pixel window), but is otherwise realistic. As mentioned in section 4.1.2, we use a single representative sky mask  $\mathbf{W}$  (and pre-filter mask  $\mathbf{W}_k$ ), and only one representative power spectrum  $\mathbf{C}$  for each of temperature and polarization. The sky mask  $\mathbf{W}$  is the average of all effective-noise-weight masks  $v_{I_i}^X$  (defined in section 4.1) in DR6; the pre-filter mask  $\mathbf{W}_k$  is the average of all pre-filter masks (one for each array) in DR6; the power spectra are defined by:

$$C_\ell = \begin{cases} (\ell/\ell_{knee})^p + 1 & \ell \geq \ell_{cap} \\ (\ell_{cap}/\ell_{knee})^p + 1 & \ell < \ell_{cap} \end{cases}, \quad (\text{B.26})$$

where for temperature  $\ell_{knee} = 3,000$ ,  $\ell_{cap} = 300$ , and  $p = -4$ , and for polarization  $\ell_{knee} = 300$ ,  $\ell_{cap} = 100$ , and  $p = -4$ . These spectra were chosen because they roughly resemble the observed noise power spectra without the Fourier filter. We measure the pseudospectra and Monte Carlo covariance matrix of the simulation ensemble. To fit for  $\alpha_{2pt}$ , we use our measured  $t_\ell$  template and optimize equation (B.21), where the mode-coupling matrix is built using the representative sky mask, against the mean and scatter of the simulated pseudospectra. To fit for  $\alpha_{4pt}$ , we use  $t_\ell$  and optimize equation (B.24) against the mean and scatter of the simulated Monte Carlo covariance matrix diagonals. Within equation (B.24), we use the INKA approximation with arithmetic symmetry (defined in section 4.1), where the mode-coupling matrix and coupling are both built using the representative sky mask. Results of those fits for the polarization case are shown in figure 11. The fit for the pseudospectra ( $\alpha_{2pt} = 0.792$ ) is excellent, with errors at the sub-percent level. The fit for the binned power spectrum covariance ( $\alpha_{4pt} = 0.469$ ) appears consistent given the noisier Monte Carlo estimates. When the final filter transfer functions —  $t_\ell^{\alpha_{2pt}}$  and  $t_\ell^{\alpha_{4pt}}$  — are applied to the fiducial signal and noise spectra in section 4.1, in cases of polarization cross-spectra, we use the geometric mean of the temperature and polarization transfer function.

We reiterate our finding that the fits are nearly independent of the mask or power spectrum used in the simulations. This is why we can use a “representative” mask and power spectra for all of DR6, and why the simulations ran here add negligible computational cost compared to the full simulation ensemble in section 4.2. Nevertheless, the procedure is ad-hoc and could be improved in the future.

### B.3 Parameters for fiducial signal spectra

The fiducial signal power spectra used in the analytic covariance matrix (section 4.1.3) and the simulations (section 4.2) follow the cosmology and likelihood foreground model of ref. [16] with model parameters given in table 3. Unlike ref. [16], we convert the frequency-dependent model components into power spectra for each frequency band in DR6 by integrating over the full array passband rather than using effective frequencies.

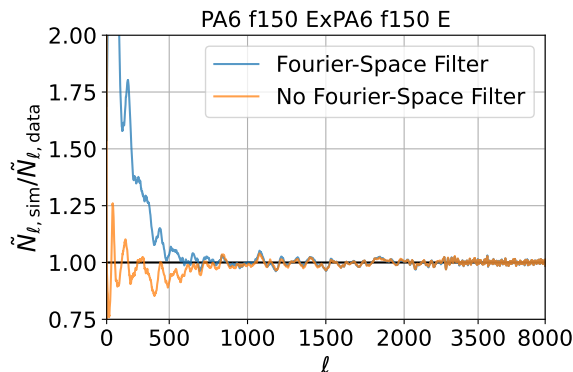
## C Updates for mnms noise simulations

In this section, we state changes to the noise model implementations of ref. [7] that were necessary for the dr6.02 maps. A more complete account of the dr6.02 noise properties and models will be made available with the release of the maps.

Parameter	Description	Value
$\log(10^{10}A_s)$	Amplitude of primordial matter power spectrum	3.044
$n_s$	Power-law index of primordial matter power spectrum	0.9649
$100\theta_{MC}$	Acoustic scale	1.04085
$\Omega_b h^2$	Physical baryon density	0.02237
$\Omega_c h^2$	Physical cold dark matter density	0.12
$\tau$	Optical depth to reionization	0.0544
$A_{tSZ}$	Amplitude of tSZ power spectrum	2.971
$A_{kSZ}$	Amplitude of kSZ power spectrum	1.6
$A_p$	Amplitude of CIB Poisson power spectrum	7.614
$\beta_p$	Spectral index of CIB Poisson power spectrum	2.2
$A_c$	Amplitude of CIB clustered power spectrum	2.755
$\beta_c$	Spectral index of CIB clustered power spectrum	2.2
$T_c$	Temperature of CIB modified blackbody	9.6 K
$\xi$	tSZ and CIB clustered correlation	0.1
$A_s$	Amplitude of point source power spectrum	3.700
$\beta_s$	Spectral index of point sources	-2.5
$A_d^{TT}$	Amplitude of dust in $TT$ power spectrum	8.83
$\alpha_d^{TT}$	Power-law index of dust in $TT$ power spectrum	-0.6
$A_d^{TE}$	Amplitude of dust in $TE$ power spectrum	0.43
$\alpha_d^{TE}$	Power-law index of dust in $TE$ power spectrum	-0.4
$A_d^{EE}$	Amplitude of dust in $EE$ power spectrum	0.165
$\alpha_d^{EE}$	Power-law index of dust in $EE$ power spectrum	-0.4
$A_d^{TB}$	Amplitude of dust in $TB$ power spectrum	0.012
$\alpha_d^{TB}$	Power-law index of dust in $TB$ power spectrum	-0.4
$A_d^{BB}$	Amplitude of dust in $BB$ power spectrum	0.116
$\alpha_d^{BB}$	Power-law index of dust in $BB$ power spectrum	-0.4
$\beta_d$	Spectral index of dust power spectrum	1.5
$T_d$	Temperature of dust modified blackbody	19.6 K

**Table 3.** Parameters for the cosmology and likelihood foreground model of ref. [16], which is specified in terms of  $D_\ell$ . The six cosmological parameters above the solid line have a flat frequency dependence. The foreground parameters below the solid line correspond to a model normalized at 150 GHz and  $\ell = 3,000$  for each component, and are rounded to three decimal places. There are no nuisance parameters in our model.

In short, there are three main changes with respect to ref. [7]. Firstly, because of a change to the map pixel window in dr6.02 [see 57], the noise at the edge of the map footprint increased compared to dr6.01. Thus, in addition to masking observed pixels and pixels with near-zero crosslinking, we also mask the 10 arcmin bordering the footprint edge, and do not include them in noise models or simulations. Secondly, as is evident in equation (3.3), the directional wavelet noise model includes the per-pixel noise standard deviations  $\sigma_i$ , whereas these were excluded in ref. [7] for this model. Lastly, we found the PA6 inverse-variance maps to contain more arcminute-scale structure compared to PA4 or PA5. When used in either



**Figure 12.** Ratio of polarization noise pseudospectra between simulations and data for PA6 f150. In the case of no Fourier-space filter being applied to the simulations or data, the ratio is consistent with unity to degree scales. Application of the Fourier-space filter to the simulations and data induces a large noise power excess in the simulations for scales larger than  $\ell \approx 500$ .

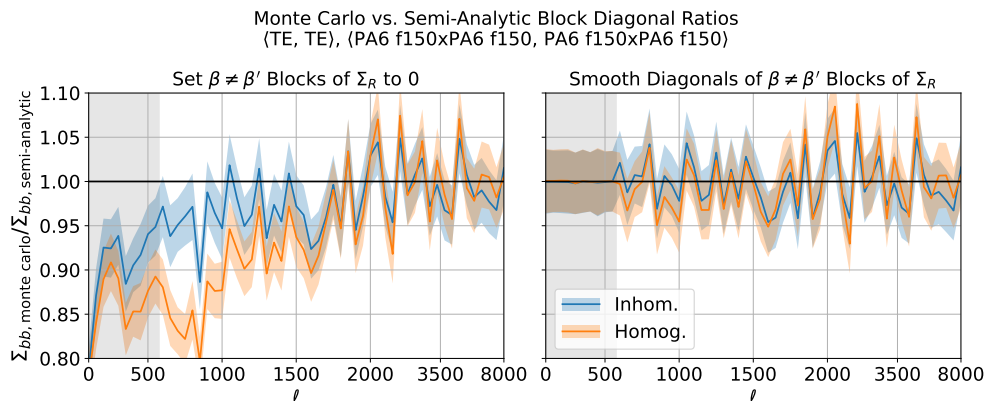
the tiled or directional wavelet noise models, these inverse-variance maps induced broadband mode-coupling that led to a percent-level but coherent noise power deficit across nearly all angular scales. We found that smoothing the PA6 inverse-variance maps with a two-arcminute Gaussian kernel eliminated the mode-coupling with no measurable unwanted side effects.

### C.1 Excess large-scale polarization noise power in simulations

The origin of the excess large scale noise power in the simulations is due to the strong noise anisotropy pattern and the sharpness of the Fourier-space filter. To see how, we start by noting that in the absence of the Fourier-space filter, the simulated noise spectra agree much more closely with the data spectra, as is apparent in figure 12. The noise power is greatest in the poorly cross-linked regions of the ACT scan strategy. As shown in figure 2, in these regions, the noise power is especially concentrated along the horizontal axis of 2D Fourier space. The Fourier-filter intentionally masks these modes: this more-optimally weights the data by cutting out a large fraction of the noise power but only a small fraction of the signal. However, this feature of the spatially-varying noise anisotropy is difficult to exactly reproduce in the simulations. If the simulations do not recover the exact noise anisotropy pattern, then after applying the filter, more noise power may persist in the unmasked modes in the simulations than in the data. In other words, application of the Fourier-space filter in this setting induces a mismatch in the per- $\ell$  average noise power between the simulations and the data where there was none before. Indeed, we find this to be the case for both noise models; figure 12 shows the result for the tiled simulations entering the baseline Monte Carlo covariance. In future work, we will optimize the noise modeling and the Fourier-space filter to enable accurate Monte Carlo covariance matrices to larger scales than those used in DR6.

## D Effect of smoothing diagonals of $\beta \neq \beta'$ blocks of $\Sigma_R$

We demonstrate the importance of including the percent-level  $\Sigma_R$  off-diagonals in the simulation-based correction of section 4.3. Specifically, we examine the bin-wise diagonals

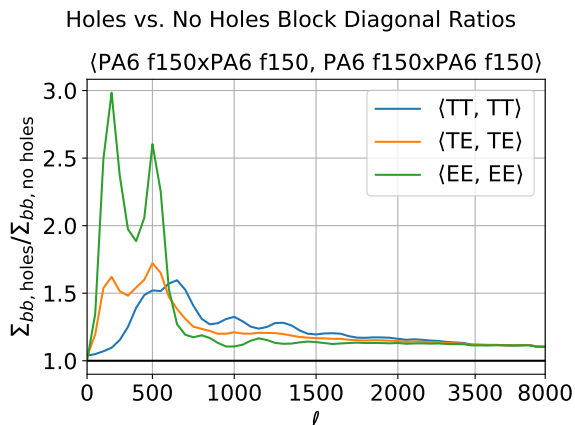


**Figure 13.** Ratios between the bin-wise diagonal of the Monte Carlo covariance matrix and the inhomogeneous and homogeneous semi-analytic covariance matrices, for the PA6 f150 x PA6 f150  $TE$  block-diagonal. *Left:* in the case that the simulation-based correction only smooths the main diagonal of  $\Sigma_R$ . *Right:* in the nominal case, where all the bin-wise diagonals of  $\Sigma_R$  are smoothed, including the off-diagonal ( $\beta' \neq \beta$ ) blocks.

( $b' = b$ ) of its off-diagonal blocks ( $\beta' \neq \beta$ ). The primary effect of neglecting these small correlations — by erroneously setting them to zero — is shown in figure 13: the amplitudes of the semi-analytic covariance bin-wise diagonals are too large compared to the Monte Carlo covariance. Notably, this has only a minor effect on the  $d_{\text{sim}}^2$  distributions of figure 9: the distributions have means of  $1757.8 \pm 2.3$  and  $1755.7 \pm 2.3$  for the inhomogeneous and homogeneous semi-analytic matrices, respectively. Both distribution means are within 0.5% of the theoretical value of 1763. Nevertheless, large biases in the covariance matrix elements themselves can be problematic for subsequent analyses that manipulate the data vector and covariance, for example, by coadding them. Fortunately, figure 13 demonstrates that accounting for these small off-diagonals in our simulation-based correction results in good agreement between the Monte Carlo and semi-analytic covariance matrices for both the inhomogeneous and homogeneous analytic prescriptions.

## E Effect of point-source holes on the covariance matrix

Although the results of section 5 suggest that the NKA does not play a significant role in the ACT DR6 covariance, it is interesting to consider its effect regardless. For instance, the literature has long recognized the questionable validity of the NKA [e.g., 11, 14, 24], especially due to analysis mask point-source holes [e.g., 12, 62]. To measure the effect of the point-source holes, we compare an analytic covariance matrix built using analysis masks that lack the holes to a matrix using the nominal masks (see figure 1). We do not apply the Fourier-space filter to the simulations or the data, and thus include no filter-correction transfer functions. This avoids two complications: as discussed in section 4.1.2, the filter-correction can inadvertently absorb biases due to the NKA, and the noise simulations do not exhibit a large power excess at low- $\ell$ . Compared to the filtered data, however, this does result in an overall steepening of the polarization noise power spectrum at low- $\ell$ . Thus, this comparison provides an upper-bound on the impact of the NKA on the DR6 covariance.

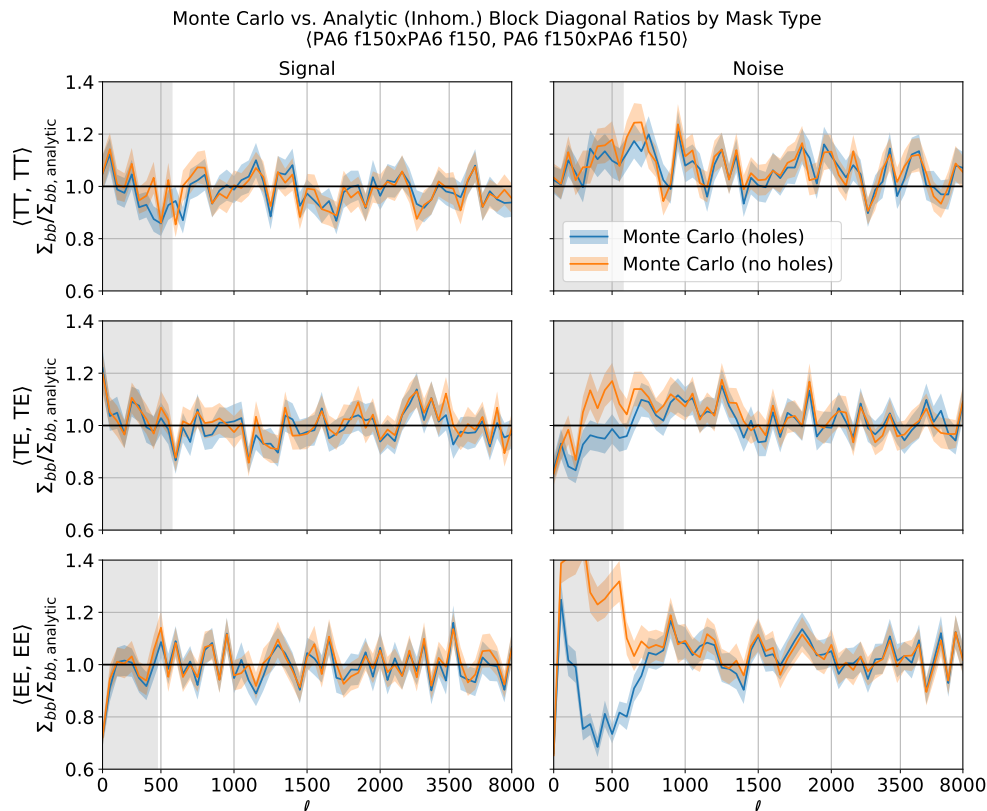


**Figure 14.** Ratios between analytic noise covariance matrix diagonals for an analysis mask including point-source holes to an analysis mask without point-source holes. For each polarization combination, the power spectrum variance increases when point-source holes are added to the mask, as expected.

We first confirm our basic intuition that, due to the broadening of the mask power spectrum, an analysis including point-source holes increases the overall covariance magnitude in figure 14. The increase is most prominent in large-scale polarization. We then assess how the performance of uncorrected analytic covariance matrices, as measured against corresponding simulations, changes due to point-source holes in figure 15. In each case — with and without point-source holes — we construct a Monte Carlo covariance from the same 600 tile-based simulations. Thus, as is evident in figure 15, both cases share the same statistical fluctuations, further facilitating a direct comparison. For the signal part of the covariance, the point-source holes induce effectively no change in the performance of the analytic matrix: the holes and no-holes cases are indistinguishable, and both exhibit percent-level agreement between simulations and the analytic matrix. As discussed in section 5.1, this contrasts with *Planck* [51, 62]; however, this is likely due to a combination of the ACT bin-width, the smaller angular scales probed, and use of the INKA for the signal rather than the traditional NKA.

The point-source holes have their largest effect on analytic matrix performance in the noise part of the covariance, especially large-scale polarization. We see the Monte Carlo covariance switches from being greater than, to being less than, the analytic covariance after adding point-source holes to the analysis mask. Neither effect is obvious in the right column of figure 8; the only difference with respect to the “holes” case here being the averaging of signal and noise, and the Fourier-filter. Thus, it is likely that a combination of the shallower polarization noise power spectra, as well as the Fourier-filter correction’s ability to absorb NKA-related biases, is operative in the baseline covariance matrix.

Interestingly, for large-scale polarization, even the no-holes case exhibits  $\sim 30\%$ -level discrepancies between the Monte Carlo and analytic matrix, in spite of its smooth analysis mask. Coupled with the fact that, whether or not we apply the Fourier-space filter, the polarization noise power spectra are shallower than either temperature noise or signal power spectra suggests that the approximate noise model of equation (3.10) is breaking down in this regime in addition to the NKA. We note that the noise anisotropy pattern in polarization is at



**Figure 15.** Ratios of Monte Carlo covariance matrix diagonals to analytic covariance matrix diagonals in the cases of an analysis mask with and without point-source holes. The format is analogous to figure 8. All results are for PA6 f150. With the exception of large-scale polarization covariance blocks, the addition of point-source holes to the analysis mask does not significantly downgrade the performance of the analytic covariance matrix.

least as strong as temperature in the poorly-crosslinked region. As we have shown in the case of the Fourier-filter, sharp anisotropies in 2D Fourier space can induce non-trivial effects at the covariance-level, regardless of their source (intrinsic to the noise, or introduced in the map processing). Further investigation is required to discern to what extent this effect could be present in the baseline covariance — that is, when including the Fourier-filter. The practical result of this study is that since future large-aperture surveys, such as SO, will contend with similar noise properties as ACT, either more work is needed to better analytically account for the effect of stripy noise properties at the covariance level, or analyses will remain reliant on Monte Carlo covariance matrices for large-scale polarization.

## References

- [1] ACT collaboration, *The Atacama Cosmology Telescope: DR4 maps and cosmological parameters*, *JCAP* **12** (2020) 047 [[arXiv:2007.07288](#)] [[INSPIRE](#)].
- [2] D. Alonso et al., *Constraining cosmology with the Gaia-unWISE quasar catalog and CMB lensing: structure growth*, *JCAP* **11** (2023) 043 [[arXiv:2306.17748](#)] [[INSPIRE](#)].

- [3] LSST DARK ENERGY SCIENCE collaboration, *A unified pseudo- $C_\ell$  framework*, *Mon. Not. Roy. Astron. Soc.* **484** (2019) 4127 [[arXiv:1809.09603](#)] [[INSPIRE](#)].
- [4] ASTROPY collaboration, *Astropy: a community python package for astronomy*, *Astron. Astrophys.* **558** (2013) A33 [[arXiv:1307.6212](#)] [[INSPIRE](#)].
- [5] ASTROPY collaboration, *The Astropy project: building an open-science project and status of the v2.0 core package*, *Astron. J.* **156** (2018) 123 [[arXiv:1801.02634](#)] [[INSPIRE](#)].
- [6] ASTROPY collaboration, *The Astropy project: sustaining and growing a community-oriented open-source project and the latest major release (v5.0) of the core package*, *Astrophys. J.* **935** (2022) 167 [[arXiv:2206.14220](#)] [[INSPIRE](#)].
- [7] Z. Atkins et al., *The Atacama Cosmology Telescope: map-based noise simulations for DR6*, *JCAP* **11** (2023) 073 [[arXiv:2303.04180](#)] [[INSPIRE](#)].
- [8] L. Balkenhol and C.L. Reichardt, *The parameter-level performance of covariance matrix conditioning in cosmic microwave background data analyses*, *Mon. Not. Roy. Astron. Soc.* **512** (2022) 4394 [[arXiv:2110.03180](#)] [[INSPIRE](#)].
- [9] SPT-3G collaboration, *Measurement of the CMB temperature power spectrum and constraints on cosmology from the SPT-3G 2018 TT, TE, and EE dataset*, *Phys. Rev. D* **108** (2023) 023510 [[arXiv:2212.05642](#)] [[INSPIRE](#)].
- [10] S. Behnel et al., *Cython: the best of both worlds*, *Comput. Sci. Eng.* **13** (2011) 31 [[INSPIRE](#)].
- [11] M.L. Brown, P.G. Castro and A.N. Taylor, *CMB temperature and polarisation pseudo- $C_\ell$  estimators and covariances*, *Mon. Not. Roy. Astron. Soc.* **360** (2005) 1262 [[astro-ph/0410394](#)] [[INSPIRE](#)].
- [12] É. Camphuis et al., *Accurate cosmic microwave background covariance matrices: exact calculation and approximations*, *Astron. Astrophys.* **668** (2022) A62 [[arXiv:2204.13721](#)] [[INSPIRE](#)].
- [13] J. Carron, *On the assumption of Gaussianity for cosmological two-point statistics and parameter dependent covariance matrices*, *Astron. Astrophys.* **551** (2013) A88 [[arXiv:1204.4724](#)] [[INSPIRE](#)].
- [14] A. Challinor and G. Chon, *Error analysis of quadratic power spectrum estimates for CMB polarization: sampling covariance*, *Mon. Not. Roy. Astron. Soc.* **360** (2005) 509 [[astro-ph/0410097](#)] [[INSPIRE](#)].
- [15] S.K. Choi et al., *Characterization of the mid-frequency arrays for advanced ACTPol*, *J. Low Temp. Phys.* **193** (2018) 267 [[arXiv:1711.04841](#)] [[INSPIRE](#)].
- [16] ACT collaboration, *The Atacama Cosmology Telescope: a measurement of the cosmic microwave background power spectra at 98 and 150 GHz*, *JCAP* **12** (2020) 045 [[arXiv:2007.07289](#)] [[INSPIRE](#)].
- [17] *HDF5 for python*, <https://www.h5py.org/>.
- [18] F. Couchot et al., *Cosmology with the cosmic microwave background temperature-polarization correlation*, *Astron. Astrophys.* **602** (2017) A41 [[arXiv:1609.09730](#)] [[INSPIRE](#)].
- [19] ACT collaboration, *Atacama Cosmology Telescope: high-resolution component-separated maps across one third of the sky*, *Phys. Rev. D* **109** (2024) 063530 [[arXiv:2307.01258](#)] [[INSPIRE](#)].
- [20] K.T. Crowley et al., *Advanced ACTPol TES device parameters and noise performance in fielded arrays*, *J. Low Temp. Phys.* **193** (2018) 328 [[arXiv:1807.07496](#)] [[INSPIRE](#)].
- [21] R. Dalal et al., *Hyper Suprime-Cam year 3 results: cosmology from cosmic shear power spectra*, *Phys. Rev. D* **108** (2023) 123519 [[arXiv:2304.00701](#)] [[INSPIRE](#)].

- [22] DES collaboration, *Dark Energy Survey year 3 results: cosmological constraints from the analysis of cosmic shear in harmonic space*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 1942 [[arXiv:2203.07128](#)] [[INSPIRE](#)].
- [23] SPT-3G collaboration, *Measurements of the E-mode polarization and temperature-E-mode correlation of the CMB from SPT-3G 2018 data*, *Phys. Rev. D* **104** (2021) 022003 [[arXiv:2101.01684](#)] [[INSPIRE](#)].
- [24] G. Efstathiou, *Myths and truths concerning estimation of power spectra*, *Mon. Not. Roy. Astron. Soc.* **349** (2004) 603 [[astro-ph/0307515](#)] [[INSPIRE](#)].
- [25] G. Efstathiou, *Hybrid estimation of CMB polarization power spectra*, *Mon. Not. Roy. Astron. Soc.* **370** (2006) 343 [[astro-ph/0601107](#)] [[INSPIRE](#)].
- [26] G. Efstathiou and S. Gratton, *A detailed description of the CamSpec likelihood pipeline and a reanalysis of the Planck high frequency maps*, *Open J. Astrophys.* **4** (2021) [[arXiv:1910.00483](#)] [[INSPIRE](#)].
- [27] DES collaboration, *Dark Energy Survey year 3 results: cosmology from galaxy clustering and galaxy-galaxy lensing in harmonic space*, *Mon. Not. Roy. Astron. Soc.* **536** (2024) 1586 [[arXiv:2406.12675](#)] [[INSPIRE](#)].
- [28] J.M. Fowler et al., *Optical design of the Atacama Cosmology Telescope and the millimeter bolometric array camera*, *Appl. Opt.* **46** (2007) 3444 [[astro-ph/0701020](#)] [[INSPIRE](#)].
- [29] DES collaboration, *Dark Energy Survey year 3 results: covariance modelling and its impact on parameter estimation and quality of fit*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 3125 [[arXiv:2012.08568](#)] [[INSPIRE](#)].
- [30] C. García-García, D. Alonso and E. Bellini, *Disconnected pseudo- $C_\ell$  covariances for projected large-scale structure data*, *JCAP* **11** (2019) 043 [[arXiv:1906.11765](#)] [[INSPIRE](#)].
- [31] C. García-García et al., *Cosmic shear with small scales: DES-Y3, KiDS-1000 and HSC-DR1*, *JCAP* **08** (2024) 024 [[arXiv:2403.13794](#)] [[INSPIRE](#)].
- [32] H. Gomersall, *pyfftw*, [Zenodo](#) (2016).
- [33] K.M. Górski et al., *HEALPix — a framework for high resolution discretization, and fast analysis of data distributed on the sphere*, *Astrophys. J.* **622** (2005) 759 [[astro-ph/0409513](#)] [[INSPIRE](#)].
- [34] B. Hadzhiyska et al., *Hefty enhancement of cosmological constraints from the DES Y1 data using a hybrid effective field theory approach to galaxy bias*, *JCAP* **09** (2021) 020 [[arXiv:2103.09820](#)] [[INSPIRE](#)].
- [35] S. Hamimeche and A. Lewis, *Properties and use of CMB power spectrum likelihoods*, *Phys. Rev. D* **79** (2009) 083012 [[arXiv:0902.0674](#)] [[INSPIRE](#)].
- [36] C.R. Harris et al., *Array programming with NumPy*, *Nature* **585** (2020) 357 [[arXiv:2006.10256](#)] [[INSPIRE](#)].
- [37] S.W. Henderson et al., *Advanced ACTPol cryogenic detector arrays and readout*, *J. Low Temp. Phys.* **184** (2016) 772 [[arXiv:1510.02809](#)] [[INSPIRE](#)].
- [38] J.C. Hill et al., *Atacama Cosmology Telescope: constraints on prerecombination early dark energy*, *Phys. Rev. D* **105** (2022) 123536 [[arXiv:2109.04451](#)] [[INSPIRE](#)].
- [39] E. Hivon et al., *Master of the cosmic microwave background anisotropy power spectrum: a fast method for statistical analysis of large and complex cosmic microwave background data sets*, *Astrophys. J.* **567** (2002) 2 [[astro-ph/0105302](#)] [[INSPIRE](#)].

- [40] S.-P.P. Ho et al., *Highly uniform 150 mm diameter multichroic polarimeter array deployed for CMB detection*, in the proceedings of the *Millimeter, submillimeter, and far-infrared detectors and instrumentation for astronomy VIII*, (2016) [DOI:10.1117/12.2233113].
- [41] H. Hotelling, *New light on the correlation coefficient and its transforms*, *J. Roy. Statist. Soc. B* **15** (1953) 193.
- [42] J.D. Hunter, *Matplotlib: a 2D graphics environment*, *Comput. Sci. Eng.* **9** (2007) 90 [INSPIRE].
- [43] L.L. Knox, *Determination of inflationary observables by cosmic microwave background anisotropy experiments*, *Phys. Rev. D* **52** (1995) 4307 [astro-ph/9504054] [INSPIRE].
- [44] WMAP collaboration, *Wilkinson Microwave Anisotropy Probe (WMAP) first year observations: TE polarization*, *Astrophys. J. Suppl.* **148** (2003) 161 [astro-ph/0302213] [INSPIRE].
- [45] C.D. Kreisch et al., *Atacama Cosmology Telescope: the persistence of neutrino self-interaction in cosmological measurements*, *Phys. Rev. D* **109** (2024) 043501 [arXiv:2207.03164] [INSPIRE].
- [46] S.K. Lam, A. Pitrou and S. Seibert, *Numba: a LLVM-based python JIT compiler*, in the proceedings of the *International conference for high performance computing, networking, storage and analysis: HPC transforms*, (2015) [DOI:10.1145/2833157.2833162] [INSPIRE].
- [47] D. Larson et al., *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: power spectra and WMAP-derived parameters*, *Astrophys. J. Suppl.* **192** (2011) 16 [arXiv:1001.4635] [INSPIRE].
- [48] O. Ledoit and M. Wolf, *Improved estimation of the covariance matrix of stock returns with an application to portfolio selection*, *J. Empirical Finance* **10** (2003) 603.
- [49] O. Ledoit and M. Wolf, *The power of (non-)linear shrinking: a review and guide to covariance matrix estimation*, *J. Financial Econometrics* **20** (2020) 187.
- [50] Y. Li et al., *Assembly and integration process of the first high density detector array for the Atacama Cosmology Telescope*, in the proceedings of the *Millimeter, submillimeter, and far-infrared detectors and instrumentation for astronomy VIII*, (2016) [DOI:10.1117/12.2233470].
- [51] Z. Li et al., *The Simons Observatory: a new open-source power spectrum pipeline applied to the Planck legacy data*, *JCAP* **09** (2023) 048 [arXiv:2112.13839] [INSPIRE].
- [52] M.J. Looijmans, M. Wang, (Shengbo) and F. Beutler, *A comparison of shrinkage estimators of the cosmological precision matrix*, *Mon. Not. Roy. Astron. Soc.* **537** (2025) 21 [arXiv:2402.13783] [INSPIRE].
- [53] T. Louis, S. Naess, X. Garrido and A. Challinor, *Fast computation of angular power spectra and covariances of high-resolution cosmic microwave background maps using the Toeplitz approximation*, *Phys. Rev. D* **102** (2020) 123538 [arXiv:2010.14344] [INSPIRE].
- [54] ACTPOL collaboration, *The Atacama Cosmology Telescope: two-season ACTPol spectra and parameters*, *JCAP* **06** (2017) 031 [arXiv:1610.02360] [INSPIRE].
- [55] P. Mahalanobis, *Reprint of: Mahalanobis, P.C. (1936) "On the generalised distance in statistics"*, *Sankhya A* **80** (2018) 1.
- [56] M. Mallaby-Kay et al., *The Atacama Cosmology Telescope: summary of DR4 and DR5 data products and data access*, *Astrophys. J. Suppl.* **255** (2021) 11 [arXiv:2103.03154] [INSPIRE].
- [57] S. Naess and T. Louis, *Large-scale power loss in ground-based CMB mapmaking*, *Open J. Astrophys.* **6** (2023) 2210.02243 [arXiv:2210.02243] [INSPIRE].

- [58] S. Naess et al., *The Atacama Cosmology Telescope: arcminute-resolution maps of 18 000 square degrees of the microwave sky from ACT 2008–2018 data combined with Planck*, *JCAP* **12** (2020) 046 [[arXiv:2007.07290](#)] [[INSPIRE](#)].
- [59] A. Nicola et al., *Cosmic shear power spectra in practice*, *JCAP* **03** (2021) 067 [[arXiv:2010.09717](#)] [[INSPIRE](#)].
- [60] G. Piccirilli et al., *Growth history and quasar bias evolution at  $z < 3$  from Quiaia*, *JCAP* **06** (2024) 012 [[arXiv:2402.05761](#)] [[INSPIRE](#)].
- [61] PLANCK collaboration, *Planck 2013 results. XV. CMB power spectra and likelihood*, *Astron. Astrophys.* **571** (2014) A15 [[arXiv:1303.5075](#)] [[INSPIRE](#)].
- [62] PLANCK collaboration, *Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters*, *Astron. Astrophys.* **594** (2016) A11 [[arXiv:1507.02704](#)] [[INSPIRE](#)].
- [63] PLANCK collaboration, *Planck 2018 results. V. CMB power spectra and likelihoods*, *Astron. Astrophys.* **641** (2020) A5 [[arXiv:1907.12875](#)] [[INSPIRE](#)].
- [64] G. Polenta et al., *Unbiased estimation of angular power spectrum*, *JCAP* **11** (2005) 001 [[astro-ph/0402428](#)] [[INSPIRE](#)].
- [65] ACT collaboration, *The Atacama Cosmology Telescope: a measurement of the DR6 CMB lensing power spectrum and its implications for structure growth*, *Astrophys. J.* **962** (2024) 112 [[arXiv:2304.05202](#)] [[INSPIRE](#)].
- [66] M. Reinecke and D.S. Seljebotn, *Libsharp — spherical harmonic transforms revisited*, *Astron. Astrophys.* **554** (2013) A112.
- [67] J. Schäfer and K. Strimmer, *A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics*, *Statist. Appl. Genetics Molecular Biol.* **4** (2005).
- [68] K.M. Surrao, O.H.E. Philcox and J.C. Hill, *Accurate estimation of angular power spectra for maps with correlated masks*, *Phys. Rev. D* **107** (2023) 083521 [[arXiv:2302.05436](#)] [[INSPIRE](#)].
- [69] R.J. Thornton et al., *The Atacama Cosmology Telescope: the polarization-sensitive ACTPol instrument*, *Astrophys. J. Suppl.* **227** (2016) 21 [[arXiv:1605.06569](#)] [[INSPIRE](#)].
- [70] M. Tristram, J.F. Macías-Pérez, C. Renault and D. Santos, *Xspect, estimation of the angular power spectrum by computing cross power spectra*, *Mon. Not. Roy. Astron. Soc.* **358** (2005) 833 [[astro-ph/0405575](#)] [[INSPIRE](#)].
- [71] EUCLID collaboration, *Euclid: covariance of weak lensing pseudo- $C_\ell$  estimates — calculation, comparison to simulations, and dependence on survey geometry*, *Astron. Astrophys.* **660** (2022) A114 [[arXiv:2112.07341](#)] [[INSPIRE](#)].
- [72] WMAP collaboration, *First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: parameter estimation methodology*, *Astrophys. J. Suppl.* **148** (2003) 195 [[astro-ph/0302218](#)] [[INSPIRE](#)].
- [73] P. Virtanen et al., *SciPy 1.0 — fundamental algorithms for scientific computing in python*, *Nature Meth.* **17** (2020) 261 [[arXiv:1907.10121](#)] [[INSPIRE](#)].
- [74] A. Zonca et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python*, *J. Open Source Softw.* **4** (2019) 1298 [[INSPIRE](#)].