Component-separated, CIB-cleaned thermal Sunyaev-Zel'dovich maps from *Planck* PR4 data with a flexible public needlet ILC pipeline

Fiona McCarthy[1,*] and J. Colin Hill[2,1,†]

[1]*Center for Computational Astrophysics, Flatiron Institute, New York, New York, USA 10010*
[2]*Department of Physics, Columbia University, 538 West 120th Street, New York, New York, USA 10027*

We use the full-mission *Planck* PR4 data to construct maps of the thermal Sunyaev–Zel'dovich effect (Compton-y parameter) in our Universe. To do so, we implement a custom needlet internal linear combination (NILC) pipeline in a Python package, PYILC, which we make publicly available. We publicly release our Compton-y maps, which we construct using various constrained ILC ("deprojection") options in order to minimize contamination from the cosmic infrared background (CIB) in the reconstructed signal. In particular, we use a moment-based deprojection that minimizes sensitivity to the assumed frequency dependence of the CIB. Our code PYILC performs needlet or harmonic ILC on mm-wave sky maps in a flexible manner, with options to deproject various components on all or some scales. We validate our maps and compare them to the official *Planck* 2015 y map, finding that we obtain consistent results on large scales and 10–20% lower noise on small scales. We expect that these maps will be useful for many auto- and cross-correlation analyses; in a companion paper, we use them to measure the tSZ—CMB lensing cross-correlation. We anticipate that PYILC will be useful both for data analysis and for pipeline validation on simulations to understand the propagation of foreground components through a full NILC pipeline.

I. INTRODUCTION

Observations of the millimeter-wave sky contain contributions from many sources, both Galactic and extragalactic. The signals of cosmological interest include, of course, the primary cosmic microwave background (CMB) [1,2]—the photons that have been free streaming since recombination without scattering—along with the Sunyaev–Zel'dovich (SZ) effect [3,4], which is sourced by the scattering of the CMB photons from free electrons in the late Universe, and the cosmic infrared background (CIB) [5–7], which is the thermal radiation of dust grains in star-forming galaxies that are heated by starlight. As we can extract very different information from the different signals, it is useful to be able to separate them. It is common to do so by observing the sky at multiple frequencies and separating the signals based on their differing frequency behavior. In particular, the CMB intensity behaves as a perfect blackbody, with a temperature of 2.726 K [8–10]. The thermal SZ (tSZ) effect—the scattering of CMB photons by high-temperature electrons—induces a well-understood distortion in this spectrum, allowing the tSZ anisotropies to be separated cleanly from the primary CMB.

Many component separation algorithms exist for performing such separation (see, e.g., [11–18]). In particular,

we focus on the internal linear combination (ILC), a method that creates linear combinations of maps using the well-understood frequency behavior of a signal of interest, such as the blackbody CMB or the tSZ effect, and which has been applied to CMB data for decades [11]. Importantly, this is a "blind" component separation technique, in that the frequency behavior of the contaminants does not need to be known *a priori*. This is necessary, as the CIB—while observed to generally behave as a modified blackbody, with its emission dominant (with respect to the CMB) at higher frequencies ($\gtrsim 353$ GHz, although on small scales the CIB anisotropies are dominant at lower frequencies than this)—does not have a well-understood spectral energy distribution (SED), or frequency dependence, that can be derived from first principles. However, when the SED of a contaminant is known, this information can be used in the ILC to create a map that is completely insensitive to this contaminant [19,20] (or partially less sensitive than the unconstrained estimate [21]), although at the expense of higher variance in the overall reconstruction. This "deprojection" of foreground components, known as a "constrained" ILC, is useful for example in the case of cross-correlations, when one is cross-correlating a component of interest with another signal that is highly correlated with a contaminant.

In this work, we use a needlet ILC (NILC) algorithm to construct maps of the tSZ effect using *Planck* PR4 (NPIPE) data [22]; we note that several y maps have

*fmccarthy@flatironinstitute.org
†jch2200@columbia.edu

previously been made with *Planck* data [23–26]. We also release maps in which we have deprojected the CIB in various ways, both with a standard CIB deprojection using an estimate of the CIB SED and in a more theory-independent method where we deproject the first moments of the CIB SED, following the methods of [27]; in a companion paper [28], we use these maps to measure the cross-correlation of the tSZ signal with CMB lensing convergence (κ) in *Planck* PR4 data using a publicly-available *Planck* κ reconstruction [29].

We validate our Compton-y maps and compare them to that from the 2015 *Planck* release [24]. While this work was in the final stages of preparation, a similar analysis performing NILC on *Planck* PR4 data appeared [26]. We defer a comparison to this map to future work.

A central purpose of this paper is also to release a user-friendly Python package PYILC, a flexible code that can perform needlet and harmonic ILC on curved-sky maps for various specifications of needlet domains. This is (to our knowledge) the first publicly available needlet ILC code.

This paper is structured as follows. In Sec. II, we describe the datasets we use to construct the ILC maps. Section III presents a pedagogical overview of the NILC algorithm and the "ILC bias." In Sec. IV, we describe our pipeline to construct a y map from the data using the NILC approach, including our exact analysis settings. Section V discusses the various CIB deprojection options that we implement. In Sec. VI, we validate our y maps, compute their power spectra, compare them to that of the official *Planck* release, and quantify the amount of CIB contamination in the various deprojections. We discuss our results and conclude in Sec. VII.

II. DATA

We use the single-frequency full-mission maps from the *Planck* NPIPE data release (PR4) [22] in our analysis. These maps are provided in µK (CMB thermodynamic temperature units) at all frequencies, so we do not need to convert between Jy/sr and µK (as might have been necessary for previous releases of the *Planck* single-frequency maps, some of which provided the highest-frequency maps in Jy/sr). Note that the kinematic solar dipole, which is sourced by the Doppler boosting of the CMB monopole due to our proper motion with respect to the CMB rest frame (in particular, our motion through the Galaxy), is not subtracted from these maps. As this is much brighter than the intrinsic CMB fluctuations (with an amplitude of ≈ 3367 µK [22], cf. the characteristic CMB fluctuation amplitude of ~ 40 µK), we subtract it from the maps before using them. To ensure we subtract the same dipole from each map, we use the solar dipole estimation from the Commander component-separation analysis of

TABLE I. Characterizing features of the *Planck* experiment, in particular, the beam FWHM and approximate white noise levels for each frequency channel. We quote the noise both in µK arcmin and in µK^2; the latter is calculated from the former by converting to µK radians [i.e., multiplying by $\pi/(180 \times 60)$] and then squaring. This information comes from Table 4 of [30]. Note that we have applied a Jy/sr-to-µK conversion factor to the values quoted for 545 and 857 GHz, as we analyze the maps in µK.

Frequency (GHz)	Beam FWHM (arcmin)	Noise (µK arcmin)	Noise power spectrum amplitude (µK^2)
30	32.29	150	0.00190
44	27.94	162	0.00222
70	13.08	210	0.00373
100	9.66	77.4	0.000507
143	7.22	33	9.21×10^{-5}
217	4.90	46.8	0.000185
353	4.92	154	0.00200
545	4.67	806.7	0.0551
857	4.22	19115	30.9

the NPIPE maps.[1] We also subtract the mean of each map before further analysis.

We use the 30, 44, 70, 100, 143, 217, 353, 545, and 857 GHz single-frequency maps. The observed sky signal in these maps is convolved with the *Planck* instrument beam at the corresponding frequency. We approximate these beams as being Gaussian, with full width at half maximum values (FWHMs) given in Table I [30]. To characterize fully the frequency response of each map, we use the passbands given in Refs. [31,32].

We also use the half-ring split maps from the NPIPE release. These maps are subset maps of the full-mission maps described above, each with the same passbands, beams, etc., as the full-mission maps described above, but with independent noise realizations; as such, each half-ring map is noisier than the full-mission map, but they are useful for power spectrum analysis as their cross-power spectrum does not contain bias due to correlations in the instrumental noise. With these maps, we build two independent split maps of the Compton-y signal.

We use a needlet ILC algorithm to construct a full-sky tSZ map from these frequency maps, with various choices for contaminant deprojection. We discuss the specific details of our needlet ILC pipeline in detail in Sec. IV.

III. THE NEEDLET ILC ALGORITHM

In this section, we discuss the ILC and NILC in general, beginning in Part III A with a general definition of ILC and describing the NILC in Part III B. We discuss the constrained ILC, which can be used to deproject specific

---

[1]This is available on NERSC at $CFS/cmb/data/planck 2020/all_data/commander_dipole_templates/ planck/dipole_CMB_n4096_K.fits.

foreground components with known frequency dependence, in Part III C. Finally in Sec. III D, we discuss the ILC bias. Note that in this section, we remain very general, specializing to the specific details of our NILC pipeline in Sec. IV.

Throughout, we use lowercase Latin indices $i, j, k, l, m, n, \ldots$ to label components of vectors or matrices in frequency space; subscript ℓ, m indices refer to discrete harmonic-space multipole coefficients (while the symbol m may seem overloaded, it will always be clear from context to what we refer). Greek indices $\alpha, \beta, \ldots$ refer to components in the space of sky components that we preserve or deproject.

A. ILC

The internal linear combination (ILC) (see, e.g., [11] for an early application reconstructing a CMB map using the *WMAP* data) relies on knowledge of the frequency dependence of a signal to isolate it by taking a linear combination of multifrequency measurements. In particular, assuming the temperature anisotropy T in frequency channel i in direction $\hat{n}$ is given by

$$T_i(\hat{n}) = a_i s(\hat{n}) + n_i(\hat{n}), \tag{1}$$

where $s(\hat{n})$ is the signal of interest, a_i is its *known* frequency dependence or spectral energy distribution (SED), and $n_i(\hat{n})$ is all other sources of intensity, which includes atmospheric or instrumental noise as well as cosmological or Galactic foregrounds (or any other signal that the detector measures). Any linear combination of the temperature maps

$$\tilde{s}(\hat{n}) = \sum_i w_i T_i(\hat{n}) \tag{2}$$

that obeys the condition

$$\sum_i w_i a_i = 1 \tag{3}$$

is unbiased to the signal of interest, in that it can be written as

$$\tilde{s}(\hat{n}) = s(\hat{n}) + \tilde{n}(\hat{n}), \tag{4}$$

where $\tilde{n}(\hat{n})$ is uncorrelated with the signal $s(\hat{n})$ (provided the noise and the foregrounds are uncorrelated with the signal—note that this assumption is in fact broken for the case of the tSZ effect, which is correlated with the cosmic infrared background and other LSS-induced foregrounds). The weights w_i in Eq. (2) that result in a minimum-variance estimate of the signal are

$$\mathbf{w} = (\mathbf{a}^T \mathcal{C}^{-1} \mathbf{a})^{-1} \mathbf{a}^T \mathcal{C}^{-1}, \tag{5}$$

where $\mathcal{C}$ is the covariance matrix of the data. These weights can be found straightforwardly by a minimization with the method of Lagrange multipliers to preserve the constraint (e.g., [33]). In explicit index notation, this can be written

$$w_i = \frac{a_j (\mathcal{C}^{-1})_{ij}}{a_k (\mathcal{C}^{-1})_{kl} a_l}, \tag{6}$$

where, here and to follow, the Einstein summation convention is assumed; i.e., repeated indices are summed over.

In the ILC approach, the frequency-frequency covariance matrix is estimated directly from the data, and so the only "external" knowledge that is needed is the knowledge of the signal's frequency dependence a_i (hence, the nomenclature "internal"). The covariance can be measured and the weights applied in various bases: for example, in real space, by measuring $\mathcal{C}_{ij}$ over the entire maps (or subregions of the maps) and calculating one weight for each frequency map or in harmonic space, by calculating an ℓ-dependent covariance matrix $\mathcal{C}_\ell^{ij}$ and applying ℓ-dependent weights to the harmonic coefficients of the frequency maps in separate multipole bins. The former is ideal if there is no scale dependence to the noise properties, e.g., if the noise and the signal all have power spectra with similar ℓ dependence. If the foregrounds have different ℓ dependence to the signal or if different foregrounds are relevant at different scales in the different frequency channels or if the instrumental noise becomes dominant at different values of ℓ, then a harmonic ILC is more appropriate. We describe briefly each of these domains below.

1. Real-space ILC

In a real-space domain $\mathcal{D}^{\text{real}}$, the temperature in frequency band i on the sphere $T_i(\Omega)$ is defined on a discrete basis of pixels $p(\hat{n})$:

$$T_i(\Omega) = \sum_{\hat{n}} T_i(\hat{n}) p(\hat{n}). \tag{7}$$

In general, the pixelized coefficients $T_i(\hat{n})$ can be found by

$$T_i(\hat{n}) = \int d\Omega T_i(\Omega) p(\hat{n}); \tag{8}$$

in the most simple version, $p(\hat{n})$ is a step function with unit value within the area defined by the pixel and zero elsewhere.

In this domain, the frequency-space covariance matrix $\mathcal{C}$ is calculated according to

$$\mathcal{C}_{ij} = \langle (T_i(\hat{n}) - \langle T_i \rangle)(T_j(\hat{n}) - \langle T_j \rangle) \rangle \tag{9}$$

$$= \frac{1}{N_{\text{pix-1}}} \sum_{\hat{n} \in \mathcal{D}^{\text{real}}} (T_i(\hat{n}) - \langle T_i \rangle)(T_j(\hat{n}) - \langle T_j \rangle), \tag{10}$$

where

$$\langle T_i \rangle \equiv \frac{1}{N_{\text{pix}}} \sum_{\hat{n} \in \mathcal{D}^{\text{real}}} T_i(\hat{n}) \qquad (11)$$

denotes the mean (over $\mathcal{D}^{\text{real}}$) of $T_i(\hat{n})$, and $N_{\text{pix}} = \sum_{\hat{n} \in \mathcal{D}^{\text{real}}} p(\hat{n})$ denotes the number of pixels in the domain $\mathcal{D}^{\text{real}}$.

If $\mathcal{D}^{\text{real}}$ covers the entire sphere (or the entire area of the unmasked map), the ILC weights then only depend on the frequency channel i and are applied uniformly to the entire map. Alternatively, $\mathcal{D}^{\text{real}}$ can be a subset of the map, allowing for spatial dependence of the weights; this may be appropriate for statistically anisotropic fields, such as when dealing with the foreground emission from our own galaxy, which is highly statistically anisotropic. However, it should be noted that, while the true temperature is preserved within the domains, correlations on scales larger than each domain are lost in such an application.

2. Harmonic-space ILC

The basis of the harmonic domain is determined by the spherical harmonic functions $Y_{\ell m}(\hat{n})$, which are the solutions to the Laplace equation on the sphere:

$$T(\hat{n}) = \sum_{\ell m} T_{\ell m} Y_{\ell m}(\hat{n}). \qquad (12)$$

The spherical harmonic coefficients $T_{\ell m}$ are related to the pixel coefficients $T(\hat{n})$ according to

$$T_{\ell m} = \sum_{\hat{n}} T(\hat{n}) Y^*_{\ell m}(\hat{n}), \qquad (13)$$

where $Y^*_{\ell m}$ denotes the complex conjugate of $Y_{\ell m}$, and the inverse transformation is

$$T(\hat{n}) = \sum_{\ell m} T_{\ell m} Y_{\ell m}(\hat{n}). \qquad (14)$$

In this domain, we can calculate the covariance matrix in an ℓ-dependent manner. Defining the measured power spectrum $\hat{C}^{ij}_\ell$ according to

$$\hat{C}^{ij}_\ell = \langle T^i_{\ell m} T^j_{\ell m} \rangle = \sum_m \frac{T^i_{\ell m} T^{j*}_{\ell m}}{2\ell + 1}, \qquad (15)$$

the covariance matrix at scale ℓ, $\mathcal{C}_{ij}(\ell)$, is given by

$$\mathcal{C}_{ij}(\ell) = \frac{2\ell + 1}{4\pi} \hat{C}^{ij}_\ell. \qquad (16)$$

In practice, it can be useful to calculate $\mathcal{C}$ in a domain $D^{\text{harm}}(\ell_0)$ defined by a *band* of multipoles, for example,

centered at some value ℓ_0 with width $\Delta\ell$. In this case, $\mathcal{C}_{ij}(\ell_0)$ is given by

$$\mathcal{C}_{ij}(\ell_0) = \sum_{\ell = \ell_0 - \Delta\ell/2}^{\ell = \ell_0 + \Delta\ell/2} \frac{2\ell + 1}{4\pi} \hat{C}^{ij}_\ell. \qquad (17)$$

The ℓ-dependent weights, calculated from Eq. (5), can then be applied to the spherical harmonic coefficients $T^i_{\ell m}$ with $\ell \in \mathcal{D}^{\text{harm}}(\ell_0)$, i.e., with $\ell_0 - \Delta\ell/2 < \ell < \ell_0 + \Delta\ell/2$. Note that it is not necessary to define disjoint bands (they can overlap, if desired); we could alternatively define a separate such bin at every ℓ, with $\mathcal{C}_{ij}(\ell_0)$ calculated using information from the surrounding multipoles.[2]

Multipole-dependent weights are appropriate when there is different scale-dependent behavior of the foregrounds and instrumental noise, including foreground SEDs that depend on ℓ, to allow for the variance to be adaptively minimized on all scales. However, the harmonic ILC, while optimal for a statistically isotropic field, is not equipped to deal with statistical anisotropy, such as that from Galactic foreground components.

B. Needlet ILC (NILC)

To combine the advantages of both the real-space and harmonic-space ILC approaches, the ILC domain can be defined on a needlet frame [15]. Needlets [34] are a construction of a spherical wavelet frame (a frame is similar to an over-complete basis), which allows for simultaneous localization in real and harmonic space. The frame is defined first by a set of harmonic-space window functions (indexed by capital Latin letters $I, J, \ldots$) h^I_ℓ, which obey

$$\sum_I (h^I_\ell)^2 = 1 \qquad (18)$$

at each ℓ. Each I specifies a different "needlet scale." The needlets are further defined by real-space domains $\mathcal{D}^{\text{real},I}_{\hat{n}}$ associated with each needlet scale, where $\mathcal{D}^{\text{real},I}_{\hat{n}}$ can be defined independently at each pixel $\hat{n}$ (and $\mathcal{D}^{\text{real},I}_{\hat{n}}$ are not disjoint over the pixels).

In practice, the needlet ILC consists of the following steps:

(1) For each needlet scale I, each frequency map i is filtered in harmonic space according to the window function h^I_ℓ:

$$T^I_i(\hat{n}) = \sum_{\ell m} h^I_\ell T^i_{\ell m} Y_{\ell m}(\hat{n}). \qquad (19)$$

---

[2]However, currently our harmonic ILC implementation in PYILC only uses disjoint ℓ bands.

The maps $T_i^I(\hat{n})$ are referred to as the "needlet coefficients." Equation (19) amounts to taking a spherical harmonic transformation of each $T_i(\hat{n})$, filtering the coefficients by multiplying them by h_ℓ^I, and taking the inverse spherical harmonic transformation of the result.

(2) For each pixel $\hat{n}$ at each needlet scale I, the local frequency-frequency covariance matrix $\mathcal{C}_{ij}(I,\hat{n})$ is calculated on a domain $\mathcal{D}_{\hat{n}}^{\text{real},I}$. This results in a set of N_I matrices of size $N_{\text{freq}} \times N_{\text{freq}}$, at each pixel $\hat{n}$:

$$\mathcal{C}_{ij}^I(\hat{n}) = \frac{1}{N_{\text{pix}} - 1} \sum_{\hat{n}' \in \mathcal{D}_{\hat{n}}^{\text{real},I}} (T_i^I(\hat{n}') - \langle T_i^I \rangle)$$
$$\times (T_j^I(\hat{n}') - \langle T_j^I \rangle), \qquad (20)$$

where N_{pix} is the number of pixels in $\mathcal{D}_{\hat{n}}^{\text{real},I}$, and $\langle T_i^I \rangle$ is the mean of the temperature on $\mathcal{D}_{\hat{n}}^{\text{real},I}$:

$$\langle T_i^I \rangle \equiv \frac{1}{N_{\text{pix}}} \sum_{\hat{n}' \in \mathcal{D}_{\hat{n}}^{\text{real},I}} T_i^I(\hat{n}'). \qquad (21)$$

(3) The needlet ILC weights $w_i^I(\hat{n})$ are then calculated at each needlet scale I according to Eq. (5). Note that the inverse of the covariance matrix $\mathcal{C}_{ij}^I(\hat{n})$ is the inverse in the frequency basis, and the inversion is performed separately at each pixel $\hat{n}$. These weights are then applied to the needlet coefficients to build the ILC estimate at each needlet scale:

$$T_{\text{ILC}}^I(\hat{n}) = \sum_i w_i^I(\hat{n}) T_i^I(\hat{n}). \qquad (22)$$

(4) The spherical harmonic coefficients of each ILC estimate are then computed and filtered (again) by the needlet window functions and then transformed back to pixel space:

$$z^I(\hat{n}) = \sum_{\ell m} h_\ell^I T_{\text{ILC},\ell m}^I Y_{\ell m}(\hat{n}). \qquad (23)$$

These final maps are then added to arrive at the final NILC estimate:

$$T^{\text{NILC}}(\hat{n}) = \sum_I z^I(\hat{n}). \qquad (24)$$

Note that the condition in Eq. (18) guarantees that signal power is preserved in this series of operations.

C. Deprojection of foregrounds: Constrained ILC

The weights in Eq. (5) are chosen to minimize the variance in the recovered map. However, sometimes certain foregrounds can significantly bias a signal of interest, and

a slight increase in variance is an acceptable price to pay for a significant reduction of this bias. This can be particularly true for cross-correlation measurements using ILC Compton-y maps, as some foregrounds are more highly correlated with the field with which we are cross-correlating than the signal we are trying to isolate. For example, the CIB is a foreground for the tSZ effect and is more highly correlated with CMB lensing than the tSZ signal is [20]. In such cases, one may choose to build a "constrained" ILC map, as described in the following.

If the SED of a foreground signal is known, we can "deproject" it from our final map. In particular, let us now say that the signal in the sky is

$$T_i(\hat{n}) = a_i s(\hat{n}) + b_i f(\hat{n}) + n_i'(\hat{n}), \qquad (25)$$

where $f(\hat{n})$ is the foreground we wish to remove and b_i is its SED. If the weights in the linear combination obey the condition

$$\sum_i b_i w_i = 0, \qquad (26)$$

the resulting signal will not contain $f(\hat{n})$. The weights that obey both Eqs. (3) and (26) and result in a minimum-variance estimate of the signal are

$$w = \mathbf{e}^T \left(\mathbf{A}^T \mathcal{C}^{-1} \mathbf{A} \right)^{-1} \mathbf{A}^T \mathcal{C}^{-1}, \qquad (27)$$

where the vector $\mathbf{e}^T = [10]$, and the matrix $\mathbf{A} = [ab]$. Written explicitly in terms of components, this is

$$w_i = \frac{(b_k(\mathcal{C}^{-1})_{kl} b_l)^{-1} (\mathcal{C}^{-1})_{ij} a_j - (a_k(\mathcal{C}^{-1})_{kl} b_l)^{-1} (\mathcal{C}^{-1})_{ij} b_j}{(a_k(\mathcal{C}^{-1})_{kl} a_l)(b_m(\mathcal{C}^{-1})_{mn} b_n) - (a_k(\mathcal{C}^{-1})_{kl} b_l)^2}. \qquad (28)$$

1. Multiply constrained ILC

This can be extended to the case where multiple components are simultaneously deprojected (see, e.g., [21,35,36]). Let us consider a case where we deproject N_{deproj} foregrounds, such that there are $1 + N_{\text{deproj}}$ constraints the ILC must obey: the signal-preserving ILC constraint for the signal of interest in Eq. (3), along with the nulling condition in Eq. (26) for the SEDs of each of the N_{deproj} components we wish to deproject. Note that there is a maximum number of foregrounds we can deproject, corresponding to $N_{\text{deproj}} = N_{\text{freq}} - 1$, as an attempt to add further constraints would overconstrain the system of equations.

In this case, the weights are given by [36]

$$
\begin{aligned}
w^i = (\mathcal{C}^{-1})_{ij} \frac{1}{\det \mathcal{Q}} &(\det(\mathcal{Q}_{1,2,\dots,N_{\mathrm{deproj}};1,2,\dots,N_{\mathrm{deproj}}})\mathcal{A}_{j0} \\
&- \det(\mathcal{Q}_{0,2,\dots,N_{\mathrm{deproj}};1,2,\dots,N_{\mathrm{deproj}}})\mathcal{A}_{j1} \\
&+ \det(\mathcal{Q}_{0,1,3,\dots,N_{\mathrm{deproj}};1,2,\dots,N_{\mathrm{deproj}}})\mathcal{A}_{j2} \\
&- \det(\mathcal{Q}_{0,1,2,4,\dots,N_{\mathrm{deproj}};1,2,\dots,N_{\mathrm{deproj}}})\mathcal{A}_{j3} + \cdots),
\end{aligned} \tag{29}
$$

where $\mathcal{Q}$ is an $(N_{\mathrm{deproj}} + 1) \times (N_{\mathrm{deproj}} + 1)$-dimensional symmetric matrix with components

$$
\mathcal{Q}_{\alpha\beta} \equiv (\mathcal{C}^{-1})_{ij}\mathcal{A}_{i\alpha}\mathcal{A}_{j\beta}, \tag{30}
$$

and where $\mathcal{A}_{i\alpha}$ is an $N_{\mathrm{freq}} \times (N_{\mathrm{deproj}} + 1)$-dimensional matrix with components given by the SEDs of the signals we wish to preserve or deproject, with the $\alpha = 0$ column that of the SED we are preserving and the $\alpha = 1, \dots$ components given by the SEDs we are deprojecting (note that the Greek indices $\alpha, \beta, \dots$ refer to these $N_{\mathrm{deproj}} + 1$ dimensions). The matrix $\mathcal{Q}_{\dots;1,2,\dots,N_{\mathrm{deproj}}}$ refers to the $N_{\mathrm{deproj}} \times N_{\mathrm{deproj}}$-dimensional submatrix of $\mathcal{Q}$, which is formed by removing the γ row and the 0th column, where γ is the index that is dropped from the first component list.

D. The ILC bias

The ILC bias is a well-known issue in the ILC algorithm [15], which is caused by the fact that the covariance matrix is estimated directly from the maps, using only a finite number of modes. Chance fluctuations lead to incorrect estimates of the covariance and can lead to correlations between the weights and the signal of interest. This can be minimized by measuring the covariance over a large enough domain with enough modes that the covariance matrix estimation is immune to chance fluctuations.

We can define the ILC bias as follows. Recall the signal estimate is

$$
\tilde{s} = \sum_i w_i T_i = s + \tilde{n}, \tag{31}
$$

where $\tilde{n}$ is given by

$$
\tilde{n} = \sum_i w_i n_i. \tag{32}
$$

The variance of $\tilde{s}$ is

$$
\langle \tilde{s}^2 \rangle = \langle (s + \tilde{n})^2 \rangle = \langle s^2 \rangle + 2\langle s\tilde{n} \rangle + \langle \tilde{n}^2 \rangle. \tag{33}
$$

The contribution from $\langle \tilde{n}^2 \rangle$ adds "noise bias" to the estimation of a power spectrum from the map and depends (to first order) on the foregrounds and the noise properties of the maps. By "ILC bias," we refer explicitly to the term

sourced by the correlation of the signal and the $\tilde{n}$ term: $\langle s\tilde{n} \rangle$. For weights constructed from the true covariance matrix of the underlying theory (and for cases where the foregrounds and the signal are uncorrelated), this term exactly vanishes. However, even in cases when the foregrounds and the signal are uncorrelated, this term can be nonzero as the weights are constructed from a covariance matrix that is *measured from the data*; thus, the signal appears in $\tilde{n}$ through w_i. We define b_{ILC} exactly as

$$
b_{\mathrm{ILC}} = \langle s\tilde{n} \rangle. \tag{34}
$$

The fractional size of the ILC bias $\frac{b_{\mathrm{ILC}}}{\langle s^2 \rangle}$ can be estimated from the number of modes used to calculate the covariance matrix, N_{modes}[3]:

$$
\frac{b_{\mathrm{ILC}}}{\langle s^2 \rangle} = \frac{|1 + N_{\mathrm{deproj}} - N_{\mathrm{freq}}|}{N_{\mathrm{modes}}}. \tag{35}
$$

Note that this bias is negative [15], but it will be convenient later in our work to have defined this quantity using an absolute value such that it is positive definite. Generically, increasing N_{modes} suppresses this bias.

In a real-space domain $\mathcal{D}^{\mathrm{real}}$, the number of modes N_{modes} is equal to the number of pixels in $\mathcal{D}^{\mathrm{real}}$. In a harmonic domain $\mathcal{D}^{\mathrm{harm}}$, the number of modes is given by $\sum_\ell (2\ell + 1)$, with the sum taken over all $\ell \in \mathcal{D}^{\mathrm{harm}}$ (here, assuming full-sky data, with the number of harmonic modes otherwise reduced by a factor of f_{sky}).

In a needlet domain $\mathcal{D}_{\hat{n}}^{\mathrm{real},I}$, the number of modes is given by $\frac{N_{\mathrm{pix}}}{N_{\mathrm{pix},\Omega}} \sum_\ell (2\ell + 1)(h_\ell^I)^2$, where N_{pix} is the (effective) number of real-space pixels in $\mathcal{D}_{\hat{n}}^{\mathrm{real},I}$; $N_{\mathrm{pix},\Omega}$ is the number of pixels on the entire sphere; and the sum is taken over all ℓ, with the needlet filter function h_ℓ^I appropriately weighting the contributions from the different multipoles. Note that this expression is only valid for real-space domains defined by top-hat window functions in pixel space; it is also possible to allow for more complicated real-space domains, such as domains defined by a real-space Gaussian window function, in which case, N_{pix} does not just count pixels but is weighted by the window function defining $\mathcal{D}_{\hat{n}}^{\mathrm{real},I}$.

IV. ESTIMATING THE tSZ SIGNAL

The tSZ spectral distortion is given by

$$
\frac{\Delta T^{\mathrm{tSZ}}(\hat{n}, \nu)}{T_{\mathrm{CMB}}} = g_\nu y(\hat{n}), \tag{36}
$$

---

[3]Note that this expression corrects an error in Eq. (3) of Ref. [24]—their numerator should be $(N_{\mathrm{channels}} - 2)$ due to the deprojection of the CMB component in their y map.

where g_ν is the tSZ spectral function [4]:

$$g_\nu = x \coth\left(\frac{x}{2}\right) - 4, \qquad (37)$$

with $x \equiv \frac{h\nu}{k_B T_{\mathrm{CMB}}}$. Here, h is Planck's constant, k_B is Boltzmann's constant, $T_{\mathrm{CMB}} = 2.726$ K is the mean temperature of the CMB [8–10], and $y(\hat{n})$ is the dimensionless (and frequency-independent) Compton y parameter that quantifies the integral of the electron pressure along the line of sight (LOS). An ILC map that preserves the SED given by Eq. (37) is a map of the Compton-y anisotropies in our Universe.

We construct a set of Compton-y maps from the single-frequency maps of the *Planck* NPIPE data release (PR4) [22]. In our "standard frequency coverage" (default) case, we use the maps at frequencies $\{30, 44, 70, 100, 143, 217, 353, 545\}$ GHz; i.e., we use all of the maps from the low frequency instrument (LFI) and all of the maps from the high frequency instrument (HFI) except for that at 857 GHz. We also make an extended-frequency-coverage version, which includes 857 GHz, and additionally a version which excludes both 545 and 857 GHz.

The NILC method has been previously applied to *Planck* data to construct y maps, e.g., in the official *Planck* analysis of the PR1 (2013) [37] and PR2 (2015) data releases [24], along with other component separation algorithms, in particular, the modified internal linear combination (MILCA) [24,25]. Reference [23] (hereafter HS14) used a harmonic ILC (HILC) to construct a Compton-y map with which to measure the tSZ-CMB lensing cross-correlation, $C_\ell^{y\kappa}$. Notably, Ref. [25] applied MILCA to the NPIPE single-frequency maps to build a lower-noise y map than the 2015 *Planck* map, and similarly Ref. [26] has recently applied NILC to the NPIPE maps.

In this section, we describe our analysis settings for the NILC algorithm. In Sec. IV A, we present the harmonic- and real-space filters we use for our NILC. We present the

preprocessing steps that we apply before performing the NILC analysis in Sec. IV B. We discuss our foreground deprojection methods subsequently in Sec. V.

A. Harmonic- and real-space filter choices for the NILC analysis

1. Harmonic-space filters

Following the official *Planck* NILC approach to construct Compton-y maps [24], we use $N_{\mathrm{scales}} = 10$ Gaussian needlet functions h_ℓ^I (where $0 < I < N_{\mathrm{scales}}$) with scales corresponding to full-width-at-half-maximum (FWHM) values of $\{600, 300, 120, 60, 30, 15, 10, 7.5, 5\}$ arcmin. To construct these needlet filters, we first define Gaussian filters G_ℓ^I for each scale $0 < I < N_{\mathrm{scales}}$ and then construct the needlet filters according to

$$h_\ell^I = \begin{cases} G_\ell^I & I = 0, \\ \sqrt{(G_\ell^I)^2 - (G_\ell^{I-1})^2} & 0 < I < N_{\mathrm{scales}} - 1, \\ \sqrt{1 - (G_\ell^I)^2} & I = N_{\mathrm{scales}} - 1. \end{cases} \qquad (38)$$

The Gaussian filters are defined as usual according to

$$G_\ell^I = \exp\left(-\frac{\ell(\ell+1)}{16 \ln 2}(\Theta_{\mathrm{FWHM}}^I)^2\right), \qquad (39)$$

where Θ_{FWHM}^I is the beam FWHM appropriate for scale I in radians. We require the needlet filters to obey the unit-transmission criterion [15]

$$\sum_I (h_\ell^I)^2 = 1 \quad \forall \; \ell. \qquad (40)$$

The Gaussian filters G_ℓ^I and the needlet filters h_ℓ^I are shown in Fig. 1.

Before performing the needlet decomposition of the single-frequency maps, we convolve them to a common
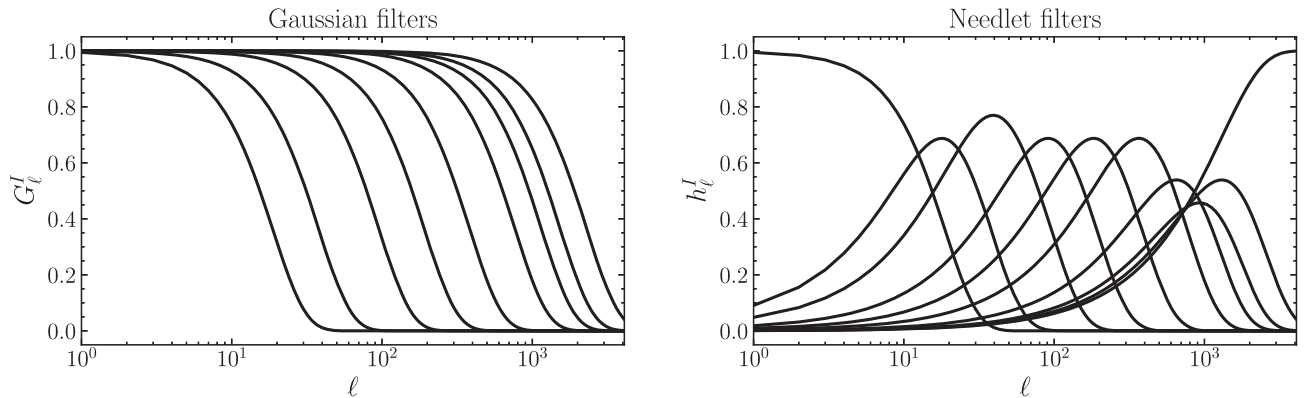

FIG. 1. The nine Gaussian filter functions (left) used to construct the ten needlet harmonic-space filters (right) according to Eq. (38). The FWHMs of the Gaussians are listed in Sec. IV A 1 (and repeated in Table II).

TABLE II.   Details of our NILC pipeline, including the FWHMs of the Gaussians used to construct the harmonic needlet filters as described in Sec. IV A 1 and the FWHMs of the real-space filters used to define the domains on which we calculate the real-space covariance matrices. The real-space domain size depends on the number of deprojected foregrounds N_{deproj}, as indicated. Note that we do not actually use the real-space filters specified here for $N_{\mathrm{deproj}} > 0$, as the larger filters calculated for $N_{\mathrm{deproj}} = 0$ will automatically satisfy the ILC bias threshold in these cases.

Needlet scale number I		0	1	2	3	4	5	6	7	8	9
Needlet scale FWHM (arcminute)		600	600-300	300-120	120-60	60-30	30-15	15-10	10-7.5	7.5-5	5
Real-space FWHMs (degrees)	$N_{\mathrm{deproj}} = 0$	373.8	216.0	81.65	43.20	20.00	9.13	6.32	4.63	2.74	0.89
	$N_{\mathrm{deproj}} = 1$	346.1	200.0	75.59	40.00	18.26	8.16	5.48	3.78	2.24	0.72
	$N_{\mathrm{deproj}} = 2$	315.9	182.6	69.01	36.51	16.33	7.07	4.47	2.67	1.58	0.51

beam. However, note that, while we perform the NILC with many input frequency maps, the lowest-frequency maps have much larger beam FWHMs than the highest-frequency maps, and the signal is exponentially subdominant to the instrumental noise on small-to-intermediate scales (after beam deconvolution). Thus, in practice, we drop particular frequency maps from the NILC on the needlet scales where we expect no information from these maps. In particular, we define a "beam threshold criterion" b_{thresh}, which we take to be $b_{\mathrm{thresh}} = 10^{-3}$. If, for a given needlet scale, the beam of a given frequency channel reaches this threshold at a lower ℓ than the needlet filter does for that scale, we drop that frequency channel from the NILC. Mathematically, we find the multipole ℓ_h, where $h^I_{\ell_h} = b_{\mathrm{thresh}}$ (on the decreasing side of the needlet filter), and the multipole ℓ_b, where $b^i_{\ell_b} = b_{\mathrm{thresh}}$. We then require that $\ell_h < \ell_b$ in order for a frequency map to be included in the NILC at this needlet scale. If this criterion is not satisfied for the beam of a given frequency channel, then that channel is dropped from the NILC at this needlet scale. Note that this procedure assumes monotonicity of the beam function and also assumes that the filter function has a peak (and thus a decreasing region). The latter assumption does not hold for the smallest-scale needlet filter (see Fig. 1), so for this scale, we simply use the same maps as used at the penultimate needlet scale.

2. Real-space filters

Our real-space domains $\mathcal{D}^{\mathrm{real},I}_{\hat{n}}$ (on which the frequency-frequency covariances are computed) are also defined by Gaussian kernels. We define their FWHM values by calculating the number of modes N_{modes} required to ensure that the fractional ILC bias [defined in Eq. (35)] is always below a tolerance b^{tol}, which we set to $b^{\mathrm{tol}} = 0.01$:

$$\frac{b_{\mathrm{ILC}}}{\langle s^2 \rangle} < b^{\mathrm{tol}}, \tag{41}$$

where $\langle s^2 \rangle$ is the variance of the signal on the domain $\mathcal{D}^{\mathrm{real},I}_{\hat{n}}$. At each needlet scale I, we convert the value of

N_{modes} implied by Eqs. (35) and (41) into FWHMs for the real-space Gaussian filters by first calculating the total number of modes on the full sky at each needlet scale I according to

$$N^I_{\mathrm{modes}} = \sum_\ell (2\ell + 1)(h^I_\ell)^2. \tag{42}$$

Noting that the fractional sky area covered by the real-space Gaussian is $2\pi\sigma^2_{\mathrm{real},I}/(4\pi) = \sigma^2_{\mathrm{real},I}/2$, we then define a variance $\sigma^2_{\mathrm{real},I}$ for our real-space filters via

$$\sigma^2_{\mathrm{real},I} = 2\left(\frac{|1 + N_{\mathrm{deproj}} - N_{\mathrm{freq}}|}{b^{\mathrm{tol}}N^I_{\mathrm{modes}}}\right), \tag{43}$$

and convert this into a FWHM with

$$\theta^2_{\mathrm{FWHM},I} = 8 \ln 2\, \sigma^2_{\mathrm{real},I}. \tag{44}$$

The resulting real-space FWHM values that we use to construct our NILC maps are listed in Table II; note that they in principle depend on the number of foregrounds deprojected N_{deproj}, with smaller real-space FWHMs

TABLE III.   The frequency maps included in our NILC algorithm at each needlet scale, given our beam threshold criterion of 10^{-3} (see Sec. IV A 1). The 857 GHz channel is listed in parentheses as it is not included in our default NILC pipeline but only in some variations thereof.

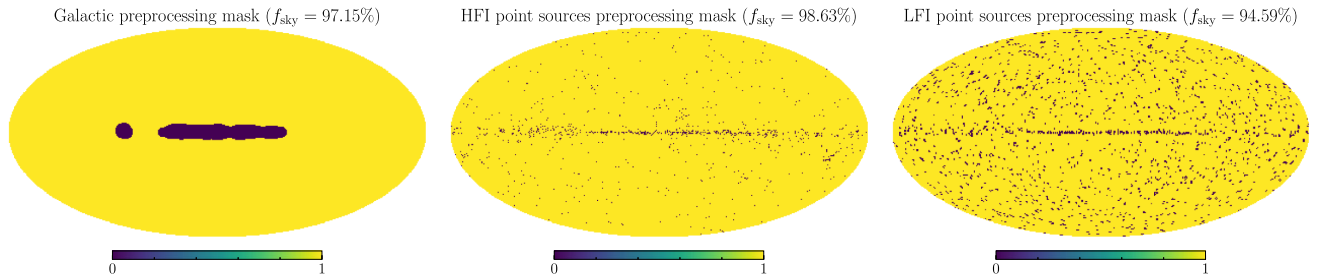| Needlet scale	Frequencies included
0	$\{30, 44, 70, 100, 143, 217, 353, 545, (857)\}$ GHz
1	$\{30, 44, 70, 100, 143, 217, 353, 545, (857)\}$ GHz
2	$\{30, 44, 70, 100, 143, 217, 353, 545, (857)\}$ GHz
3	$\{30, 44, 70, 100, 143, 217, 353, 545, (857)\}$ GHz
4	$\{44, 70, 100, 143, 217, 353, 545, (857)\}$ GHz
5	$\{70, 100, 143, 217, 353, 545, (857)\}$ GHz
6	$\{100, 143, 217, 353, 545, (857)\}$ GHz
7	$\{143, 217, 353, 545, (857)\}$ GHz
8	$\{143, 217, 353, 545, (857)\}$ GHz
9	$\{143, 217, 353, 545, (857)\}$ GHz

FIG. 2.   The masks used in the preprocessing of our needlet ILC pipeline. These masks are applied to the single-frequency maps, which are then inpainted by iteratively replacing each masked pixel with the mean of the surrounding unmasked pixels (neighbors for an unmasked pixel) before we run the needlet ILC algorithm.

allowed when deprojecting more components (at fixed b^{tol}). However, in practice, we compute the covariance matrices only *once*, using the filters of the $N_{\mathrm{deproj}} = 0$ case, as the covariance matrix calculation and inversion is the most computationally intensive part of the NILC. We then save the covariance matrices and inverse covariance matrices for use in the remaining NILC calculations for different deprojection choices. As the real-space domains used in the $N_{\mathrm{deproj}} = 0$ case are larger than those required to satisfy the ILC bias threshold for $N_{\mathrm{deproj}} > 0$, the ILC bias threshold will be satisfied in all cases. Note that, in practice, N_{freq} depends on the needlet scale, as we drop the lower-resolution maps from the higher-resolution needlet scales according to Table III.

Once we have defined the domains $\mathcal{D}_{\hat{n}}^{\mathrm{real},I}$, we calculate the mean and the frequency-frequency covariance of the needlet coefficients $T_i^I(\hat{n})$ on them by smoothing each quantity with a Gaussian beam of the appropriate FWHM. In particular, the mean is simply given by the smoothed needlet coefficient maps, while the covariance is calculated by subtracting these means from the full needlet coefficient maps and multiplying them together, then smoothing the result.

B. Preprocessing details

Before applying the NILC algorithm to the data, we apply a mask to each single-frequency map to remove very bright Galactic emission and point sources. This prevents the dominant emission from these regions (which, even if retained, would be masked in the final analysis) from significantly affecting the ILC weights calculated slightly away from (but in the vicinity of) these regions and thus allows for better component separation away from these regions.

We inpaint the regions of each map covered by these masks using a diffusive inpainting scheme, which iteratively replaces masked pixels with ≥ 4 unmasked neighbors by the mean of the surrounding unmasked pixels. The Galactic mask we use is the same as that used in the preprocessing of the *Planck* 2015 NILC tSZ map [24], which we download from the Planck Legacy Archive

(PLA).[4] The point source mask we use is the point source catalog mask from *Planck*, which we also download from the PLA, using the low-frequency catalog for the LFI maps and the high-frequency catalog for the HFI maps.[5] These preprocessing masks are shown in Fig. 2. The Galactic mask covers 2.85% of the sky; the HFI point source mask covers 1.37% of the sky; and the LFI point source mask covers 5.41% of the sky. The combination of the Galactic mask and the HFI point source mask covers 3.93% of the sky and the combination of the Galactic mask and the LFI point source mask covers 7.94% of the sky; the combination of all three covers 8.50% of the sky, which at the end, defines the total masked sky area of our final NILC maps.

As the NPIPE single-frequency maps retain the dominant contribution from the kinematic dipole, before inpainting, we remove the kinematic dipole as measured by the Commander component separation algorithm [22] and then subtract the remaining monopoles of the maps. Finally, following Ref. [24], we deconvolve the beams of each map (the beams of the *Planck* maps are listed in Ref. [30]; we repeat them in Table I) and reconvolve all maps to a common beam of $10'$. Note that, as the covariances are calculated in real space, the beam at which the NILC is performed can significantly affect the final result, as the needlet filters are so broad in scale that the large variance of the small-scale modes can contribute significantly if they are not sufficiently beam convolved. Thus, the weights can spuriously adapt to mitigate the variance sourced by the smallest-scale modes in these frequency channels. Such an impact could be mitigated by choosing needlet scales with less broad coverage in ℓ space. In this work, however, we choose to follow the official *Planck* analysis and perform the NILC on maps at $10'$ resolution.

---

[4]This is available at http://pla.esac.esa.int/pla/aio/product-action?MAP.MAP_ID=COM_CompMap_Compton-SZMap-nilc-ymaps_2048_R2.00.fits (field = 3).

[5]These are available at http://pla.esac.esa.int/pla/aio/product-action?MAP.MAP_ID=HFI_Mask_PointSrc_2048_R2.00.fits (HFI) and http://pla.esac.esa.int/pla/aio/product-action?MAP.MAP_ID=LFI_Mask_PointSrc_2048_R2.00.fits (LFI).

V. DEPROJECTION OF VARIOUS COMPONENTS

In Sec. III C, we discussed constrained ILC and foreground deprojection. In this section, we present our specific choices for foreground deprojection, in particular, for removing the CMB and the CIB.

A. Deprojection of the CIB

For many cross-correlation analyses of the tSZ signal with large-scale structure (LSS) tracers, it is necessary to deproject the CIB, which is also correlated with LSS and which can bias a $\langle y \times LSS \rangle$ measurement if unmitigated. However, unlike the tSZ and CMB signals, which display no frequency decorrelation and which have SEDs that are well understood from first principles and can thus be calculated theoretically, the CIB is not described perfectly by one SED. It is sourced by the line-of-sight integrated thermal emission of different objects; in particular, different frequency channels are sensitive to slightly different objects, as source emission at different redshifts will be redshifted into different frequency bands. This leads to frequency decorrelation between the CIB channels. However, as the correlation coefficients are $\gtrsim 90\%$ at the frequencies of interest [38–41], it is still possible to clean the CIB using multifrequency measurements. Its SED does not need to be known for the unconstrained ILC; however, if we wish to explicitly deproject it, we must model its SED.

Let us explicitly write the CIB intensity as

$$I_\nu^{\mathrm{CIB}}(\hat{n}) = \Theta_\nu^{\mathrm{CIB}} A^{\mathrm{CIB}}(\hat{n}), \qquad (45)$$

where all frequency dependence is absorbed in the CIB SED, $\Theta_\nu^{\mathrm{CIB}}$, and $A^{\mathrm{CIB}}(\hat{n})$ is a template of the CIB anisotropies. Here, we model the CIB as a modified blackbody, with an SED given by

$$\Theta_\nu^{\mathrm{CIB}} \propto \nu^\beta B_\nu(T_{\mathrm{CIB}}^{\mathrm{eff}}), \qquad (46)$$

where $B_\nu(T)$ is the Planck function

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}. \qquad (47)$$

The SED depends on two parameters: the spectral index β and the effective CIB temperature $T_{\mathrm{CIB}}^{\mathrm{eff}}$. We stress that $T_{\mathrm{CIB}}^{\mathrm{eff}}$ is not a physical temperature, but a parameter in this effective description. Also note that the overall normalization of the CIB SED in Eq. (46) is arbitrary.[6] For concreteness, we will thus write the CIB intensity as

---

[6]One only needs to specify the overall normalization for the component that is preserved in an ILC map so that the output map is in the correct units; the components that are deprojected do not need their normalization specified.

$$I_\nu^{\mathrm{CIB}}(\hat{n}) = \Theta_\nu^{\mathrm{CIB}} A^{\mathrm{CIB}}(\hat{n}) = \left(\frac{\nu}{\nu_0}\right)^{\beta+3} \frac{1}{e^{x_{\mathrm{CIB}}} - 1} A^{\mathrm{CIB}}(\hat{n}), \quad (48)$$

where $A^{\mathrm{CIB}}(\hat{n})$ is the underlying CIB component template, and with $x_{\mathrm{CIB}} \equiv \frac{h\nu}{k_B T_{\mathrm{CIB}}^{\mathrm{eff}}}$. We introduce a pivot frequency ν_0 in Eq. (48), which one can choose arbitrarily; we set $\nu_0 = 353$ GHz. It is important to note that this quantifies the CIB intensity, not temperature; natural units for the CIB intensity are Jy/sr or MJy/sr. As we analyze our maps in units of CMB thermodynamic temperature μK, we must convert from intensity to μK in order to write the SED of the CIB as is relevant for our maps. This is a frequency-dependent conversion, which is obtained by differentiating the Planck function [Eq. (47)] with respect to T and is given by

$$\left(\frac{dB_\nu}{dT}\right)_{T=T_{\mathrm{CMB}}} = \frac{2h\nu^3}{c^2} \frac{e^x}{(e^x - 1)^2} \frac{x}{T_{\mathrm{CMB}}}, \qquad (49)$$

where $x \equiv \frac{h\nu}{k_B T_{\mathrm{CMB}}}$.

The parameters of the effective CIB SED ($\beta, T_{\mathrm{CIB}}^{\mathrm{eff}}$) can be fit to observations of the CIB. For use in our fiducial CIB deprojections, we use a determination of the CIB SED determined by a fit to the monopole predictions of the best-fit halo model of Ref. [39], in particular, their Table 10. We reproduce these predictions for the CIB SED in Table IV.

We write a simple Gaussian likelihood for the data in Table IV with 1σ errors given by the quoted error bars and no covariance between the different frequencies:

$$-2 \ln \mathcal{L}(\beta, T_{\mathrm{CIB}}^{\mathrm{eff}}, A) = \sum_\nu \frac{(A\Theta_\nu^{\mathrm{CIB}}(\beta, T_{\mathrm{CIB}}^{\mathrm{eff}}) - d_\nu)^2}{\sigma_\nu^2}, \quad (50)$$

where d_ν are the data points in Table IV (with the units appropriately converted), σ_ν are the 1σ errors reported in Table IV, and $\Theta_\nu^{\mathrm{CIB}}(\beta, T_{\mathrm{CIB}}^{\mathrm{eff}})$ is the quantity defined in Eq. (46). Note that the expression $A\Theta_\nu^{\mathrm{CIB}}$ is the monopole analog of Eq. (45).

The parameters that maximize the likelihood are found by straightforwardly minimizing the right-hand side of Eq. (50). When we only include information from the three lower frequencies $\nu = \{217, 353, 545\}$ GHz in the likelihood, the maximum-likelihood parameters are given by $\beta = 1.77$, $T_{\mathrm{CIB}}^{\mathrm{eff}} = 10.14$ K; when we additionally include 857 GHz, the best-fit parameters are $\beta = 1.59$, $T_{\mathrm{CIB}}^{\mathrm{eff}} = 11.95$ K. To quantify the uncertainty, we perform

TABLE IV. The predicted values of the CIB monopole in Ref. [39], as calculated from the halo model fit to the CIB power spectra in that work at $\nu = \{217, 353, 545, 857\}$ GHz.

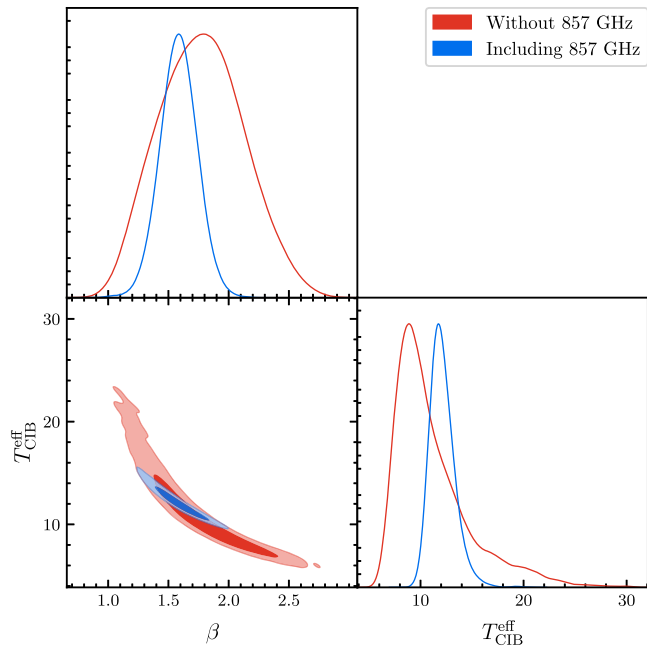| Frequency [GHz]	217	353	545	857
$\nu I_\nu [\mathrm{nW/m^2/sr}]$	0.077 ± 0.003	0.53 ± 0.02	2.3 ± 0.1	7.7 ± 0.2

FIG. 3. The posterior for the parameters of the inferred effective CIB SED. There is significant degeneracy between the parameters β and $T_{\rm CIB}^{\rm eff}$. For the without-857-GHz case, the likelihood is maximized at $\beta = 1.77$, $T_{\rm CIB}^{\rm eff} = 10.14$ K.

Markov-chain Monte Carlo (MCMC) sampling of the posterior, using flat, linear priors on all parameters, resulting in a posterior region shown in Fig. 3. For this, we use the Metropolis-Hastings algorithm implementation of [42,43] implemented in Cobaya[7] [44,45]. We run the chains until they are converged with a Gelman-Rubin criterion [46] of $|R - 1| < 0.005$. The marginalized means and posteriors of β and $T_{\rm CIB}^{\rm eff}$ are shown in Table V, along with the best-fit values.

We stress that these are not intended to be highly accurate posteriors for the CIB SED parameters, as we have not quantified appropriately the covariance of the data points. Instead, they are intended to give an idea of a realistic region of variation of the parameters. Note that for the case when we do not include 857 GHz, we are fitting three data points with three parameters, and so our reduced χ^2 is ill defined, and we do not have a proper notion of goodness of fit (and, as expected, the χ^2 approaches 0). For the case when we do include 857 GHz, we find that the model is a good fit, as quantified by the reduced χ^2.

The best-fit SEDs are plotted in Fig. 4, along with the SED previously used in Ref. [47] to deproject the CIB, which used $T_{\rm CIB}^{\rm eff} = 24$ K and $\beta = 1.2$. The latter SED slightly overpredicts the 545 GHz emission and severely overpredicts the 857 GHz emission, although, note that the highest frequency channel used in Ref. [47] was the *Planck* 545 GHz channel (857 GHz was not used). Finally, note

---

[7]https://cobaya.readthedocs.io/en/latest/.

that using these (approximate) CIB SEDs requires extrapolation to lower frequencies (<217 GHz) where the SED is not directly constrained.

1. A note on the best-fit β, $T_{\rm CIB}^{\rm eff}$ and the halo model parameters

Often, the CIB emission is modeled with a halo model with certain parameters that can be fit to data. Many of these halo models model the SED of objects at a given redshift as a modified blackbody and include in the model a functional form for the (possibly z-dependent) physical dust temperature T_D and spectral index β_D (see, e.g., [48,49]). Often $T_D(z)$ is parametrized as follows:

$$T_D(z) = T_0(1+z)^\alpha, \tag{51}$$

with T_0 and α free parameters that are fitted in the analysis. For example, the halo model analysis of [39] found $T_0 = 24.4 \pm 1.9$ K and $\alpha = 0.36 \pm 0.05$, along with $\beta_D = 1.75 \pm 0.06$.

If one wants to directly use these parameters to deproject the CIB sourced at a given redshift, it is important to account for the appropriate redshifting of the CIB temperature. For example, to deproject the CIB emission from $z = 1$, one can calculate the temperature of the modified blackbody at $z = 1$ as 31.31 K and then redshift it to today: $T(z)/(1+z) = 15.66$ K (β_D is unaffected by the redshifting operation and can be directly used in the CIB SED). For $\alpha = 1$ (as for the evolution of the CMB temperature), this would be equivalent to always using T_0. The CIB monopole is sourced at $z \approx 2$, which would predict a redshifted temperature of ≈ 12.1 K, not far from our best-fit effective temperatures.

B. Deprojection of the first moments of the CIB

There is significant uncertainty on the parameters (β, $T_{\rm CIB}^{\rm eff}$) of the CIB SED; additionally, it is not a perfect approximation to the CIB emission to describe it as an exact modified blackbody described by a single set of these parameters. For cross-correlations of the tSZ effect with tracers that are more highly correlated with the CIB than with the tSZ field itself (e.g., the CMB lensing potential), imperfect removal of the CIB due to a deprojection with the "incorrect" SED can result in a CIB residual that is still a significant bias to the measurement. This systematic must therefore be considered and mitigated when measuring such a cross-correlation, such as in our companion paper [28] where we measure the tSZ-CMB lensing cross-correlation.

In Ref. [27], a moment-based approach for the description and removal of mm-wave foregrounds was introduced. In particular, their method was designed to mitigate effects due to SED uncertainty or variation of SED parameters due to either spatial averaging within an instrument's beam or along the line of sight; the latter is particularly important for

TABLE V.   The marginalized means and best-fit values of the β, $T_{\mathrm{CIB}}^{\mathrm{eff}}$ parameters. We also indicate the reduced χ^2 by explicitly writing the calculated χ^2 and the number of degrees of freedom, which is equal to the number of data points minus the number of fitted parameters (recall that we fit an amplitude along with β and $T_{\mathrm{eff}}^{\mathrm{CIB}}$).

| | Without 857 GHz | | With 857 GHz | |
	Mean	Best fit	Mean	Best fit
β	$1.75^{+0.39}_{-0.37}$	1.77	1.60 ± 0.16	1.59
$T_{\mathrm{CIB}}^{\mathrm{eff}}$	$11.87^{+0.75}_{-4.94}$ K	10.14 K	$12.01^{+0.99}_{-1.38}$ K	11.95 K
χ^2/d.o.f.		$1.53 \times 10^{-15}/(3-3)$		$0.302/(4-3)$

foregrounds such as the CIB, which is not emitted at only one redshift but in fact has a wide redshift kernel and is a superposition of imperfect modified blackbodies at various redshifts.

In this approach, the underlying (fundamental) SED is parametrized by some free parameters **p**, such as a modified blackbody parametrized by (β, T). One then considers variations of these parameters within the instrument beam (and/or sky map pixel) or along the line of sight, both of which generically lead to deviations of the "measured" SED away from the fundamental SED form. The observed SED can be represented as a Taylor expansion (in the parameters **p**) around a point in parameter space $\bar{\mathbf{p}}$ and more accurately modeled using this Taylor expansion. The zeroth-order moment of the observed SED is the average SED, which can be written as the fundamental SED evaluated at $\bar{\mathbf{p}}$. If there were no variations in the SED parameters, then no further modeling would be needed. However, in general, such variations lead to higher-order terms (moments) in the Taylor expansion. Then, the more appropriate SED to use for deprojection in a constrained ILC would be not only the SED evaluated at $\bar{\mathbf{p}}$ but the entire Taylor expansion. In particular, the SED and its first (and higher-order) derivatives with respect to **p** (evaluated at $\bar{\mathbf{p}}$)

can be considered as independent components to be deprojected in a constrained ILC. Of course, given a *finite* number of frequency channels, we do not have the freedom to deproject an infinite number of components, but with the multifrequency coverage of *Planck*, it is possible to deproject several components.

We will use this approach to deproject the CIB in a robust manner from our y map. In particular, we deproject the first moment of the CIB with respect to β, which in practice, amounts to deprojecting an additional component with an SED given by $\frac{\partial I_\nu^{\mathrm{CIB}}}{\partial \beta}$, and also the first moment of the CIB with respect to $T_{\mathrm{CIB}}^{\mathrm{eff}}$, which similarly amounts to deprojecting an additional component with an SED given by $\frac{\partial I_\nu^{\mathrm{CIB}}}{\partial T_{\mathrm{CIB}}^{\mathrm{eff}}}$.

These moments are given explicitly by

$$\frac{\partial I_\nu^{\mathrm{CIB}}(\hat{n})}{\partial \beta} = \ln\left(\frac{\nu}{\nu_0}\right) I_\nu^{\mathrm{CIB}}(\hat{n}), \tag{52}$$

$$\frac{\partial I_\nu^{\mathrm{CIB}}(\hat{n})}{\partial T_{\mathrm{CIB}}^{\mathrm{eff}}} = I_\nu^{\mathrm{CIB}}(\hat{n}) \frac{x_{\mathrm{CIB}}}{T_{\mathrm{CIB}}^{\mathrm{eff}}} \frac{e^{x_{\mathrm{CIB}}}}{e^{x_{\mathrm{CIB}}} - 1}. \tag{53}$$
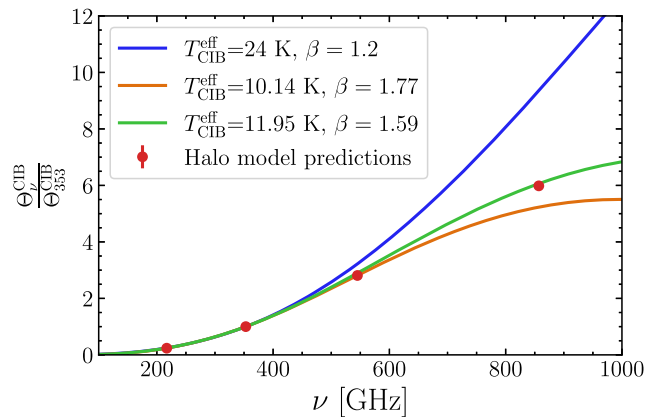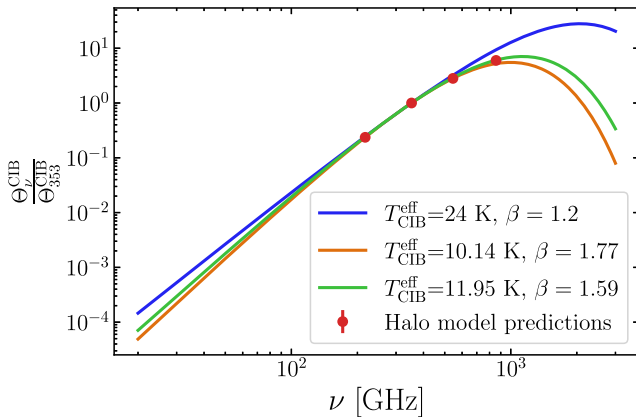

FIG. 4.   The best-fit CIB SEDs, both including the 857 GHz data [green line, $(T_{\mathrm{CIB}}^{\mathrm{eff}}, \beta) = (11.95 \text{ K}, 1.59)$] in the fit and excluding it [orange line, $(T_{\mathrm{CIB}}^{\mathrm{eff}}, \beta) = (10.14 \text{ K}, 1.77)$]. We also include for comparison the SED used in Ref. [47] [$(T_{\mathrm{CIB}}^{\mathrm{eff}}, \beta) = (24 \text{ K}, 1.2)$], which slightly overpredicts the 545 GHz emission and severely overpredicts the 857 GHz emission, but which is compatible with the data at the $1 - 2\sigma$ level (see the posteriors in Fig. 3). The left and right panels show the same data and theory curves but with a logarithmic (linear) y axis on the left (right).

For concision, we sometimes refer to these moment components as $\delta\beta$ and $\delta T_{\mathrm{CIB}}^{\mathrm{eff}}$, respectively. Note that, unlike the SED of the CIB itself, the moment expansion of β depends explicitly on the chosen pivot frequency ν_0, as at this frequency, the CIB SED does not depend on β and thus the derivative exactly vanishes.[8] In all cases, we use a pivot frequency of $\nu_0 = 353$ GHz.

We refer the interested reader to Ref. [28] for details of how this moment deprojection technique stabilizes the measurement of $C_\ell^{y\kappa}$ even for different choices of the CIB SED parameters $(\beta, T_{\mathrm{CIB}}^{\mathrm{eff}})$.

C. Deprojection of the CMB

The CMB, whose SED is known (nearly) perfectly, can be deprojected in a constrained ILC tSZ map with no uncertainty or need to consider additional moments. This can be useful for large-scale cross-correlations of the tSZ signal with LSS, as the large-scale CMB contains contributions sourced by LSS arising from the integrated Sachs-Wolfe (ISW) effect [50]. This ISW-LSS correlation (e.g., [51–53]) can bias such measurements, including the tSZ—CMB lensing cross-correlation. However, we do not have enough frequency channels to deproject more than three components at small scales (unless 857 GHz information is included); see Table III. Thus, we elect to only deproject the CMB from the first five needlet scales in some cases, denoted CMB[5]. As the ISW is a large-scale effect, this should mitigate essentially all of the LSS bias due to the ISW contribution.

The (normalized) SEDs of various components of interest are plotted in Fig. 5.

VI. COMPTON-y MAP VALIDATION AND COMPARISON TO *PLANCK* MAP

In this section, we present and validate our Compton-y maps, estimate their power spectra, and compare them to the official *Planck* 2015 tSZ NILC map. We explicitly compute the autopower spectra and 1D histograms of our maps and compare them to those of the official *Planck* analysis (on the same region of sky) in Sec. VI B. In Sec. VI C, we quantify the variance increase resulting from our various deprojections by comparing the power spectra of our undeprojected and deprojected maps; we also illustrate the effects of varying the frequency coverage in the NILC. Finally, in Sec. VI D, we compare the level of CIB contamination in our maps to that of the official *Planck* release and also present the results of deprojecting the CIB.
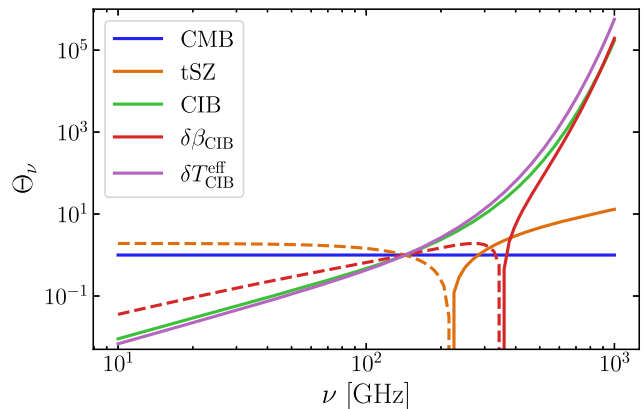

FIG. 5. The SEDs of various components of interest: the tSZ effect (orange), which we want to isolate; the CIB (green), which we wish to deproject; and the first moments of the CIB with respect to β (red) and $T_{\mathrm{CIB}}^{\mathrm{eff}}$, which we also wish to deproject. The SEDs are shown in thermodynamic temperature units in which the blackbody CMB is constant, as indicated by the CMB SED in blue. Due to the logarithmic y axis, we show negative values with dashed lines. We normalize all of the SEDs so that their absolute value is unity at 143 GHz. In all cases, we use the best-fit CIB SED parameters of $\beta = 1.77$, $T_{\mathrm{CIB}}^{\mathrm{eff}} = 10.14$ K.

We perform all power spectrum calculations throughout with `NaMaster` [54].

Note that the 2015 *Planck* map was made with the CMB deprojected, and so all direct comparisons between our map and the *Planck* map are done with our CMB-deprojected map.

A. Map images

We show our resulting tSZ maps in Fig. 6, including various deprojection choices and also showing the official *Planck* 2015 NILC tSZ map [24] for comparison. For the cases when just the CIB or the CIB + CMB are deprojected, there is no significant change to the maps by eye. However, when we deproject $\delta\beta$, the map becomes visibly noisier; we see that when we deproject both $\delta\beta$ and the CMB, the map is even noisier, indicating that there is significant CMB contamination in the CIB + $\delta\beta$-deprojected map. Finally, the maps with both moments of the CIB deprojected have significantly higher variance, and it becomes difficult to pick out structures such as the Coma cluster, which appears as a bright point in the northern hemisphere (right hand side of the maps) by eye. In the legend of Fig. 6, and later, CMB[5]-deprojection refers to the deprojection of the CMB on only the first five needlet scales.

B. Comparison with *Planck* map: Histograms and power spectra

To compare directly with the *Planck* y map, we consider histograms of the Compton-y values in the maps and also compute their autopower spectra (for our various

---

[8]When we deproject components in a constrained NILC, the normalization of their SED does not matter (when we preserve a component in the NILC, its normalization is important in order to interpret correctly the units of the final map). In the case of the SED of the CIB, ν_0 appears as a normalization and can be divided out. The same is not true for the first moment with respect to β, as ν_0 does not appear multiplicatively.
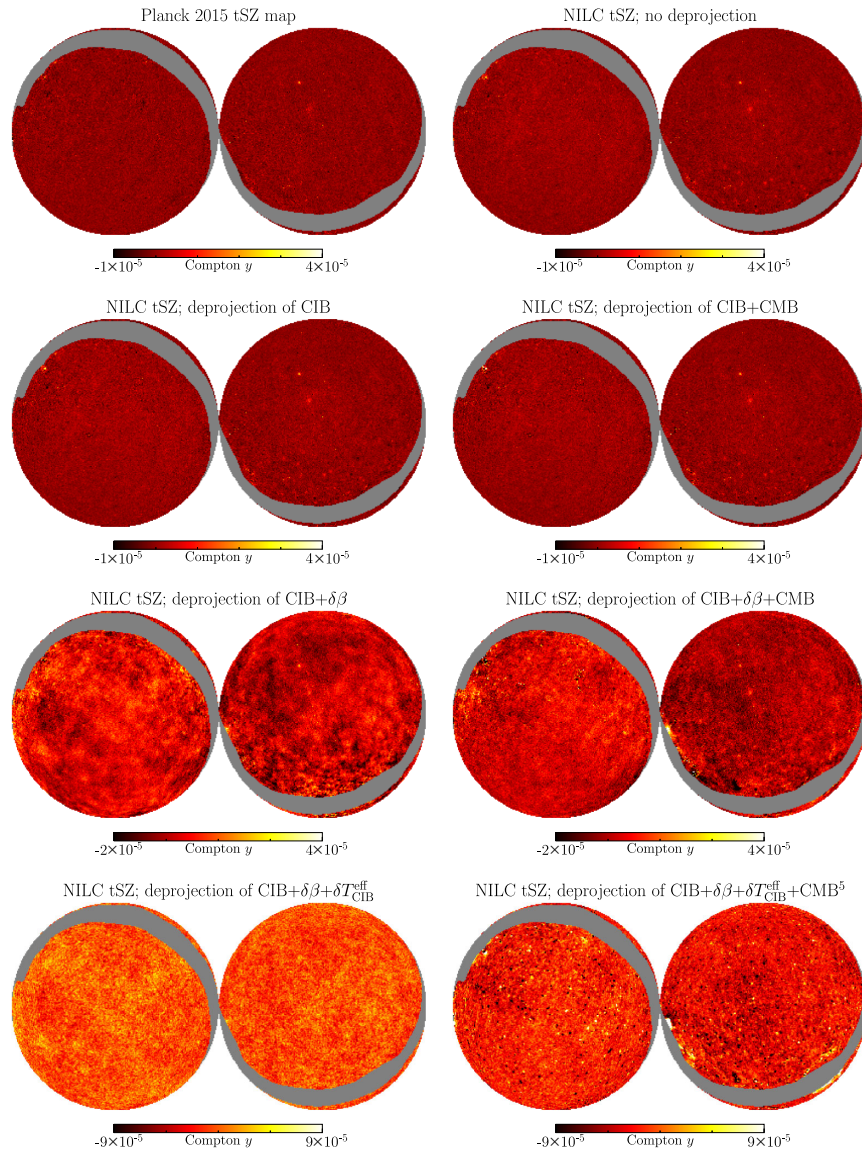
FIG. 6. Our needlet ILC maps, visualized in orthographic projection in equatorial coordinates. In all cases, the northern hemisphere is on the right and the southern on the left. We also show, on the top left, the official *Planck* 2015 NILC tSZ map [24]. In all cases, we have masked out the Galaxy with the *Planck* Galactic plane mask, which covers 20% of the sky. Note that for these visualizations, we show Gaussian-beam-convolved maps, which have FWHM = 10′. We note that the *Planck* y map has deprojected the CMB; we do not show our CMB-deprojected map here, but it is indistinguishable by eye from the no-deprojection, CIB-deprojected, and CIB+CMB-deprojected maps. Indeed, it is only when we deproject $\delta\beta$ that adding the CMB deprojection makes a visible difference (as seen in the third row). Note the increased ranges on the color bars in the bottom four plots due to their significantly increased variances.

deprojection choices) on the area of sky defined by the *Planck* y map analysis mask, which is provided with the *Planck* y map.[9] This mask, which is apodized, covers 50% of the sky. We multiply this mask by the point source mask, defined by the union of the *Planck* HFI and LFI point source masks as described in Ref. [24]. When we estimate the power spectra, we apodize the point source mask with an apodization scale of 30′ before multiplying it with the *Planck* analysis mask (for the histograms, we use an unapodized point source mask). The resulting apodized mask allows for analysis with an effective $f_{\rm sky} = 0.45$.[10]

---

[9]It can be found at http://pla.esac.esa.int/pla/aio/product-action?MAP. MAP_ID=COM_CompMap_Compton-SZMap-masks_2048_R2.01.fits, with `field=1`.

---

[10]Note that, at some level, the point source mask is correlated with the tSZ field itself; however, any biases resulting from this correlation (e.g., [55–57]) are expected to be negligible at *Planck* sensitivity.
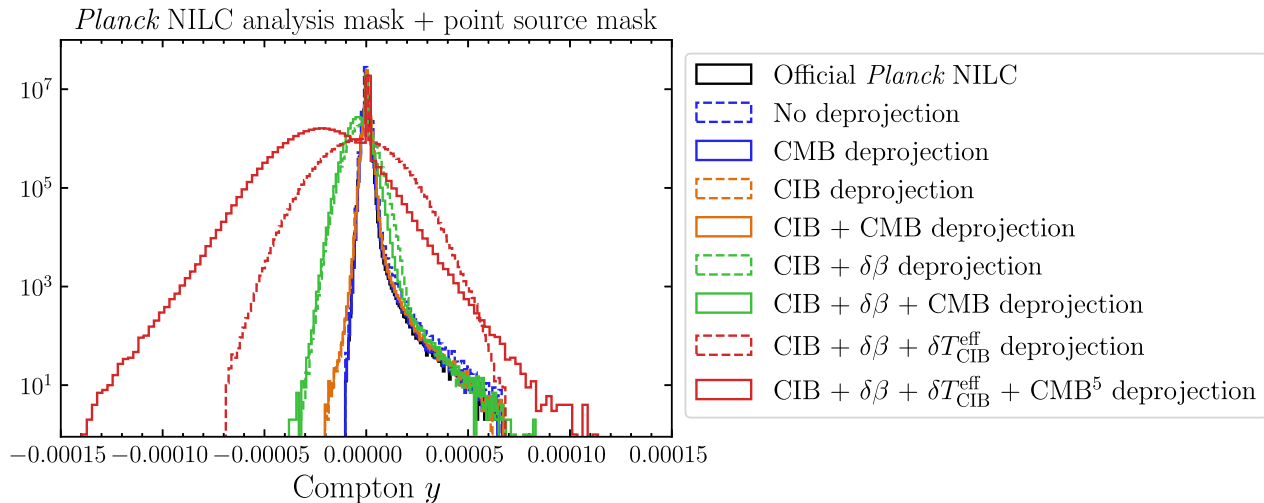
FIG. 7.   Histograms of the Compton-y values of our tSZ maps (for various deprojection choices as labeled), as compared to that of the official *Planck* 2015 tSZ map. These histograms are estimates of the one-point probability density function (PDF) of the tSZ effect (which itself is a sensitive cosmological probe [58,59]). We show the histograms of the maps convolved with a FWHM = 10′ beam and apply the *Planck* NILC analysis mask and the point source mask described in the text. The positive tails in the histograms indicate the presence of tSZ signal, with the increased noise in the CIB-deprojected and CIB + $\delta\beta$ maps evident as an increased spread around the peak at $y = 0$. The histogram of our no-deprojection and our CMB-deprojected maps agree very well with each other and with that of the official *Planck* map, demonstrating how small the signal-to-noise penalty is for deprojecting the CMB; the CIB-deprojected versions are slightly noisier (note that the CIB+CMB-deprojected version is not meaningfully noisier than the CIB-deprojected version, with the orange dashed and solid lines lying on top of each other). Once we deproject the moments of the CIB, there is a larger increase in noise, although for the $\delta\beta$-deprojected version, the non-Gaussian positive tail is still visible. When we deproject both moments, the contribution from foregrounds is larger than this tail, and it cannot be seen by eye; we also see that for the case when both moments are deprojected, there is a significant noise penalty when we also deproject the CMB (note that we do not deproject the CMB on all scales, as we do not have enough frequency coverage; instead, we just do so on large scales, in particular, in the first five needlet scales, referred to as "CMB⁵-deprojection" in the legend).

1. Histograms

We show the histograms of the Compton-y pixel values in Fig. 7 for our maps with no deprojection (standard ILC, blue), fiducial CIB modified blackbody SED deprojected (orange), and additionally with the first moment of the CIB SED with respect to β deprojected (green) and with both moments deprojected (red); we include all versions both with and without the CMB deprojected, with the CMB-deprojected versions shown in solid lines compared to the dashed lines for the cases when we do not deproject the CMB (we only deproject the CMB on the first five needlet scales in the case where we deproject both moments of the CIB due to the insufficient frequency coverage on small scales to deproject more components). We also show the histogram for the official *Planck* 2015 tSZ map in black. The Compton-y signal is highly non-Gaussian, and the strong positive tail is indicative of the presence of tSZ signal due to groups and clusters, which we see in all of the maps except for the case when we deproject both moments of the CIB; in this case, the large foreground contribution is larger than this signal. We note that there is more noise in the CIB-deprojected maps and significantly more noise in the moments-deprojected maps, as indicated by the broadening of the histograms. This is an inevitable consequence of deprojecting more components in a constrained ILC.

We note that the histograms of our no-deprojection (standard ILC) and CMB-deprojected y maps agree very well with that of the *Planck* map.

2. Autopower spectra

We measure the power spectra of our y maps to further characterize their properties. We also construct NILC y maps with, and measure the cross-power spectra of, two independent split maps that have independent noise (and therefore no noise bias in the power spectrum). We show the results in Fig. 8. We also calculate the autopower spectrum of the *Planck* y map on the same sky area. In all cases, we calculate the power spectra using `NaMaster` to decouple the mask mode-coupling matrix.

In Fig. 8, and in all subsequent plots with autopower spectra, we include a theoretical calculation of the tSZ power spectrum, calculated using CLASS_SZ[11] [60–62], which is an extension of the cosmological Boltzmann solver CLASS[12] [63]. This signal is computed in the halo model using the pressure-mass relation of Ref. [64].

[11]https://github.com/borisbolliet/class_sz.
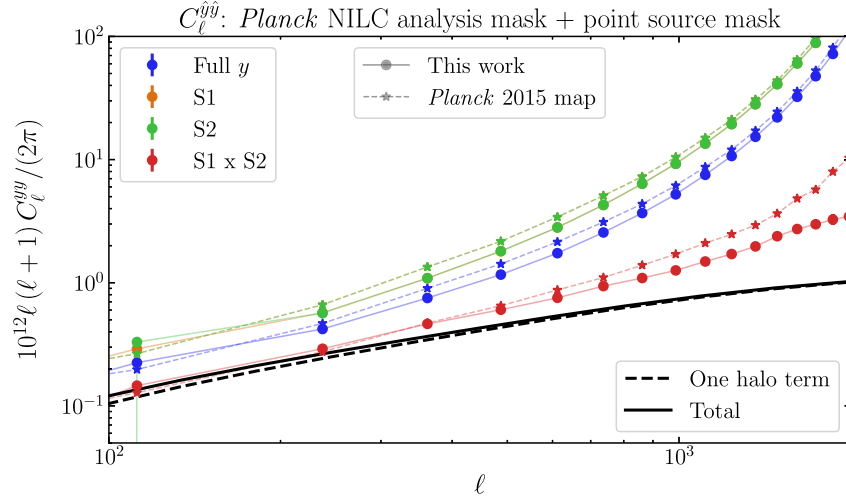[12]https://lesgourg.github.io/class_public/class.html.

FIG. 8. The autopower spectrum of our Compton-y maps, estimated by taking the auto-power of the full-mission y map (blue points) and also of the split maps (orange and green points, which nearly coincide), and the cross-power spectrum of the split maps (red points). We show the power spectra estimated both from our maps (circles) and from the 2015 *Planck* tSZ map (stars). Our maps show a clear reduction in noise power, as assessed from the autospectra and also a reduction in the residual foreground power, as assessed from the cross-spectra, compared to the 2015 *Planck* release. For reference, we also show a theoretical prediction of the tSZ power spectrum in black (total signal in solid, one-halo term only in dashed).

We refer the reader to [28] for a detailed discussion of the modeling of this signal.

In general, the power spectrum error bars can be estimated with the Gaussian expression for the covariance:

$$\mathbb{C}(\hat{C}_\ell^{\alpha\beta}, \hat{C}_{\ell'}^{\gamma\delta}) = \frac{\delta_{\ell\ell'}}{(2\ell+1)f_{\text{sky}}}((C_\ell^{\alpha\gamma} + N_\ell^{\alpha\gamma})(C_\ell^{\beta\delta} + N_\ell^{\beta\delta})$$
$$+ (C_\ell^{\alpha\delta} + N_\ell^{\alpha\delta})(C_\ell^{\beta\gamma} + N_\ell^{\beta\gamma})), \quad (54)$$

where $C_\ell^{\alpha\beta}$ and $N_\ell^{\alpha\beta}$ indicate the signal and noise including all sources of foregrounds; in practice, we replace these with the measured power $\hat{C}_\ell^{\alpha\gamma}$ such that

$$\sigma^2(\hat{C}_\ell^{yy}) = \frac{2(\hat{C}_\ell^{yy})^2}{(2\ell+1)f_{\text{sky}}}. \quad (55)$$

We do not calculate Eq. (54) directly but instead use `NaMaster`, which computes the Gaussian covariance accounting properly for the decoupling of the mask. Note that this requires us to have an estimate of $C_\ell + N_\ell$ at every ℓ, which we achieve by unbinning our measured $\hat{C}_\ell$.

The autopower spectrum of our standard ILC y map is slightly lower ($\approx$10–20%) on small scales than that of the *Planck* 2015 y map, presumably due to the lower instrumental noise in the PR4 `NPIPE` data compared to PR3. In addition, we also see a significant decrease in the cross-power spectrum of our split maps compared to that of the *Planck* 2015 y map splits, except for on the largest scales. This is presumably due to the improved foreground cleaning in our map, resulting from the lower noise in the `NPIPE` maps.

We defer a full analysis of the tSZ autopower spectrum from our maps to future work, as this will require careful understanding of the propagation of foreground contaminants, likely using simulations (as in Ref. [24]).

C. Effect of various deprojections and varying frequency coverage

In Fig. 9, we show how various contaminant deprojection choices used in our NILC analyses affect the final autopower spectrum of the y map. In all cases, when we deproject the CIB and its first moments, we use the best-fit CIB SED parameters from Sec. VA, $\beta = 1.77$ and $T_{\text{CIB}}^{\text{eff}} = 10.14$ K. We see that deprojecting the CIB incurs a small increase in noise power, but the CMB deprojection leaves the power essentially unaffected, as long as either no other components or the CIB alone is deprojected. This is because the CMB is already the dominant astrophysical contaminant in the standard ILC, so the weights are quite similar regardless of whether it is deprojected. However, deprojecting the moments of the CIB incurs much more serious noise penalties, particularly if multiple moments are deprojected. Note that, for the case when we deproject both moments of the CIB, there are not enough frequency channels to also deproject the CMB on small scales, and so we only deproject the CMB in the first five needlet scales; we refer to this in the legends of Figs. 9–11 as CMB[5].

Our fiducial y map includes information from all *Planck* frequency channels except for 857 GHz. For comparison, we also construct maps including 857 GHz in the NILC, as well as dropping 545 GHz. We note that the official *Planck* analysis (as well as the analysis of Ref. [26]) uses 857 GHz
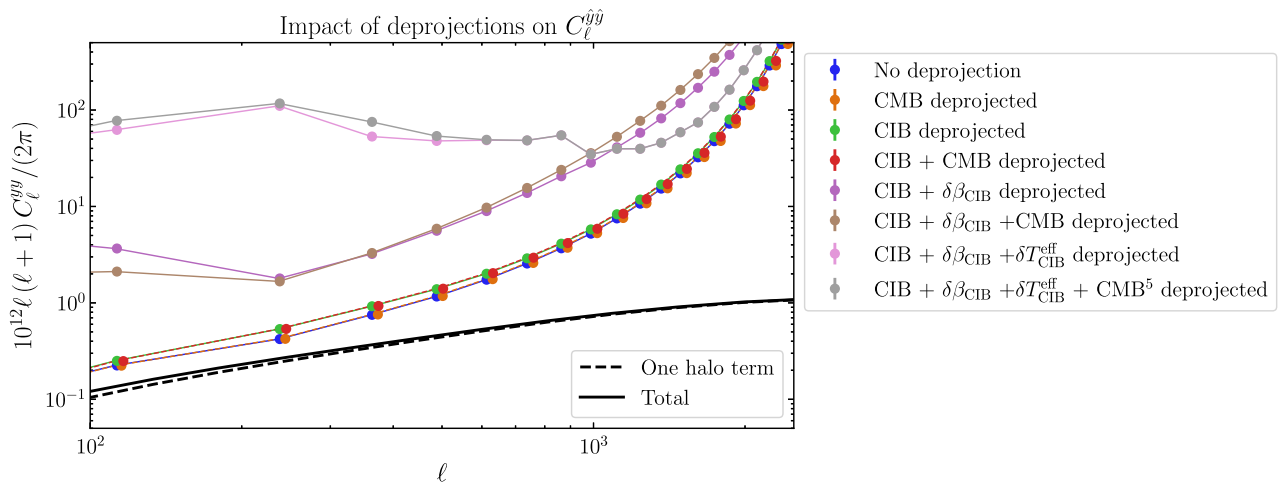
FIG. 9.   Autopower spectra of our tSZ maps for different deprojection choices, on the region of sky defined by the *Planck* analysis mask and our point source mask. Note that, to aid in distinguishability, we have offset horizontally the orange and red CMB-deprojected and CIB+CMB-deprojected points slightly, as otherwise they would fall almost exactly on the blue and green points, respectively. We have also connected these points with a dashed instead of a solid line so that the lines underneath are visible. We see that the first CMB deprojection is almost free in terms of its noise penalty; for CIB deprojection, in contrast, there is a slight noise penalty, corresponding to an increase of $\approx 10\%$ in the power (although an additional CMB deprojection here continues to increase the power negligibly). Deprojecting the first moments of the CIB, however, incurs a much larger noise penalty, with the power increasing by a factor of ≈ 5 compared to the no-deprojection power when we include $\delta\beta$ in the deprojection. Additionally, deprojecting the CMB is no longer "free" in this case, with the brown points slightly higher than the purple points. Finally, deprojecting both $\delta\beta$ and $\delta T_{\mathrm{CIB}}^{\mathrm{eff}}$ incurs a serious penalty, with the power increasing by orders of magnitude at low ℓ, although surprisingly decreasing at high ℓ compared to the CIB $+ \delta\beta$ deprojection. As there are not enough frequency channels to simultaneously deproject all of these components *and* the CMB at high resolution, we only do so for the first five needlet scales (grey points); this again incurs a small penalty compared to the case without deprojecting the CMB.
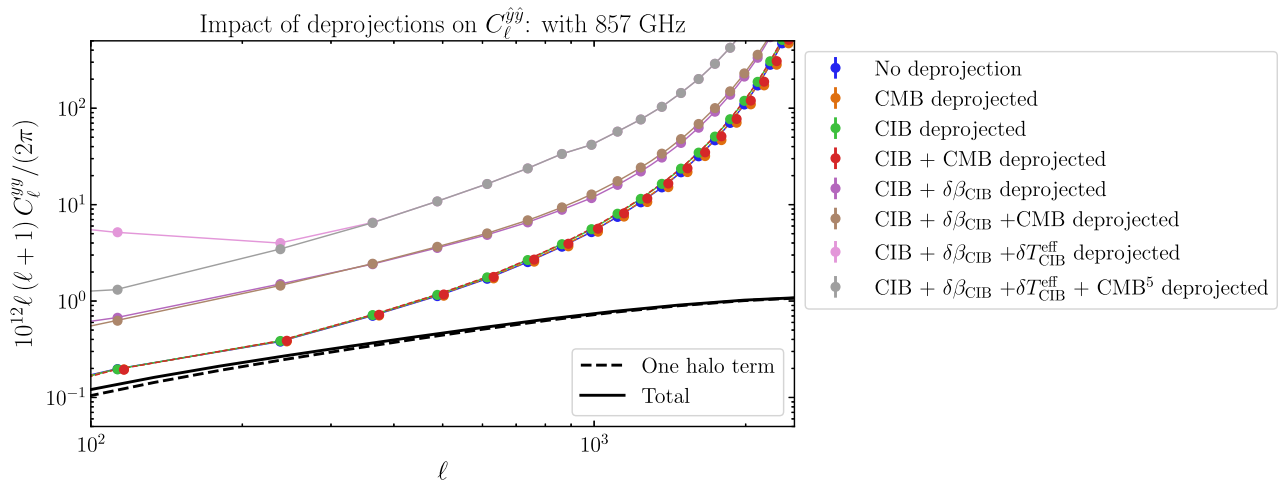

FIG. 10.   The effect of different deprojections on a y map constructed using a NILC that includes the 857 GHz channel. The noise penalty incurred when deprojecting many moments of the CIB is lower here than in the case when 857 GHz is not included in the NILC (see Fig. 9). Again, we have slightly offset the CMB-deprojected (orange) points from the undeprojected (blue) points and the CIB+CMB-deprojected (red) points from the CIB-deprojected (green) points.

only at $\ell < 300$, due to its uncertain calibration. In principle, one could co-add the $\ell < 300$ modes of our with-857 GHz map with the $\ell > 300$ modes of our no-857 GHz map to obtain such a map.

The effects of including 857 GHz and dropping 545 GHz are shown in Figs. 10 and 11, respectively. As expected, as there is more (less) data, including more (fewer) frequency channels results in lower (higher) power in the final y map.
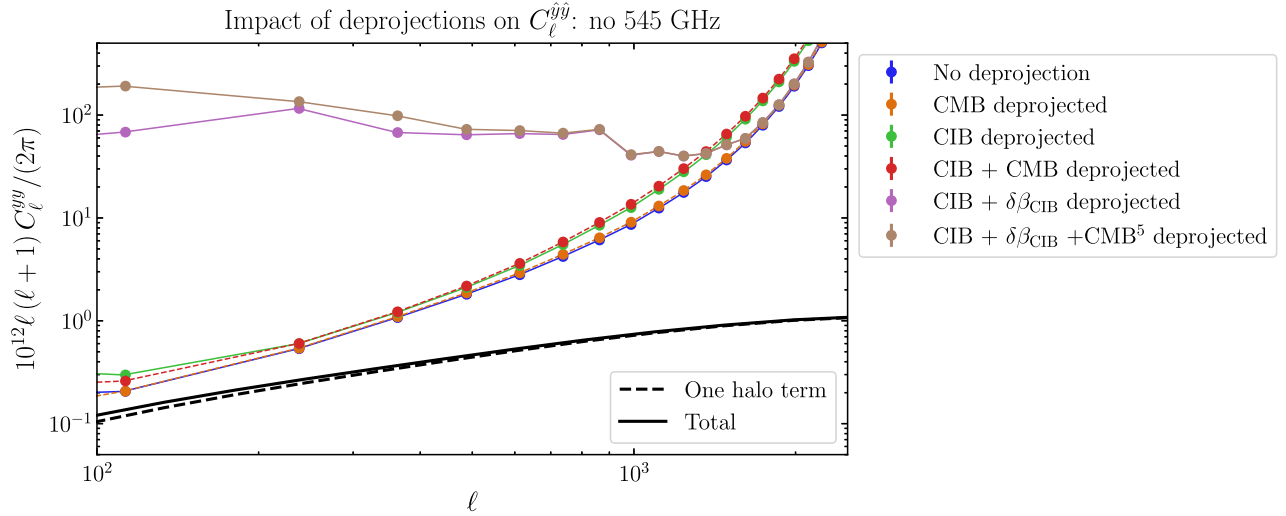
FIG. 11. The effect of different deprojections on a y map constructed using a NILC that excludes the 545 GHz channel (857 remains excluded as well). There are not enough frequency channels to simultaneously deproject the CIB along with $\delta\beta$ and $\delta T_{\mathrm{CIB}}^{\mathrm{eff}}$ in this case. Note that, in contrast to Fig. 9, we have not offset the orange CMB-deprojected points or the red CIB+CMB-deprojected points as they are more distinguishable by eye here than they are in Fig. 9 due to the slightly increased noise penalty when deprojecting the CMB.

However, we note that when more high-frequency channels are included, more information about the CIB SED is required in order to deproject it robustly, and it can be more difficult to robustly remove all CIB contamination; we discuss this issue in detail in our companion paper [28].

D. CIB contamination

We attempt to quantify the amount of residual CIB in our Compton-y maps by directly cross-correlating them with maps of the CIB. CIB maps are difficult to construct due to the difficulty of separating CIB emission from Galactic dust emission, which is dominant on large scales and has a very similar frequency dependence to the CIB. We use the CIB maps of Ref. [41], which were constructed by using HI data [65] as a tracer of the Galactic dust in order to clean the dust emission from the highest-frequency (353, 545, 857 GHz) *Planck* channels. The resulting maps of the CIB emission at these frequencies are very clean of Galactic contamination.

For this investigation, we use only the 857 GHz CIB map, as the maps were constructed with the PR3 *Planck* data and thus do not have maps built from the same independent noise realizations as the NPIPE maps. Thus, we do not have two independent splits that are required to make a measurement without instrumental noise bias. However, as we do not use 857 GHz in our default NILC y map construction, we can make a measurement with the 857 GHz map that is immune to noise bias. It would be interesting in the future to repeat or extend the work of Ref. [41] to the NPIPE data in order to measure cross-correlations with the lower frequencies that do not contain noise bias. An additional advantage of using the 857 GHz map for the tSZ-CIB cross-correlation is that this is the channel in which the intrinsic y signal is most

suppressed compared to the CIB. For lower frequencies (particularly 353 GHz and below), we would have to consider the $\langle y^{\mathrm{CIB-map}} y^{\mathrm{NILC-map}} \rangle$ contribution to the signal.

We show in Fig. 12 the cross-power spectra measured with the 857 GHz CIB map and our standard-frequency-coverage NILC y maps for several deprojection choices. We also include the measurement using the official *Planck* map. We perform these measurements with the cleanest map provided by Ref. [41], which covers 8.71% of the sky. We multiply the apodized mask provided by Ref. [41] with the mask used previously throughout this work, i.e., the official *Planck* NILC mask combined with the point source masks. The final sky coverage is 8.47% in our tSZ-CIB cross-correlation measurement.

In Fig. 12, it is clear that the CIB contamination in our y maps is lower than that of the official *Planck* y map. This is consistent with the results of Ref. [26]. It is encouraging that the lower instrumental noise in the NPIPE maps has allowed for improved subtraction of the foregrounds in this case.

In Fig. 12, we also include a theoretical prediction for the intrinsic cross-power spectrum between the tSZ and the CIB emission at 857 GHz C_ℓ^{y857}. We note that there is significant model uncertainty in this prediction. The methodology that we use follows the halo model formalism used to calculate the autopower spectrum. While we defer to our companion paper [28] an in-depth presentation of the halo model details, we note that the model in Fig. 12 (which was calculated with CLASS_SZ) uses the pressure profiles of Refs. [66,67] to predict the tSZ signal along with the halo model of Ref. [48] with the parameter values of Ref. [39] to predict the CIB signal (see Ref. [49] for a detailed description of this incarnation of the CIB halo model). As
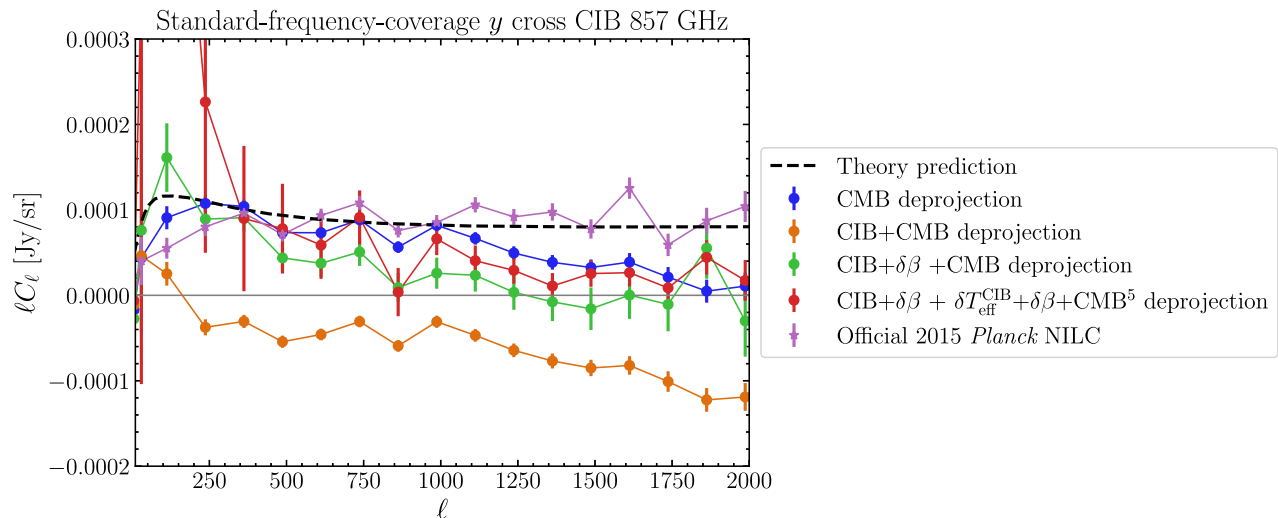
FIG. 12.   The cross-correlation of our Compton-y maps with the 857 GHz CIB map of Ref. [41] for different deprojection choices, along with the measurement for the official *Planck* NILC y map (in black). We also show a prediction for the intrinsic C_ℓ^{y857} signal, calculated in CLASS_SZ using the halo model (details in the main text). We note that the cross-correlation of our CMB-deprojected map with the CIB is significantly lower than that of the official *Planck* map; a similar conclusion was found in Ref. [26].

the focus of this paper is not the CIB-tSZ cross-correlation, we do not consider a wider range of models (although note that a prediction for this signal was made in Ref. [68], and it would be interesting to use this data to constrain this model). We note that, if it is truly CIB-free, the fully moment-deprojected $(\mathrm{CIB} + \delta\beta + \delta T_{\mathrm{eff}}^{\mathrm{CIB}} + \mathrm{CMB}^5)$ data points in Fig. 12 can be interpreted directly as a measurement of the intrinsic cross-correlation of Compton-y and the CIB. Such a measurement was previously made with

Planck data in Ref. [69]. We defer to future work a comparison between our data and the measurements in this reference. Indeed, our work in Ref. [28] relies on the assumption that the fully moment-deprojected y map is indeed negligibly contaminated by CIB. We note that the $\mathrm{CIB} + \delta\beta + \delta T_{\mathrm{eff}}^{\mathrm{CIB}} + \mathrm{CMB}^5$-deprojected points in Fig. 12 are indeed stable to variations in the assumed CIB SED, as shown in Fig. 13 (the same is not true of the CIB deprojection or the $\mathrm{CIB} + \delta\beta$ deprojection). In Fig. 12,
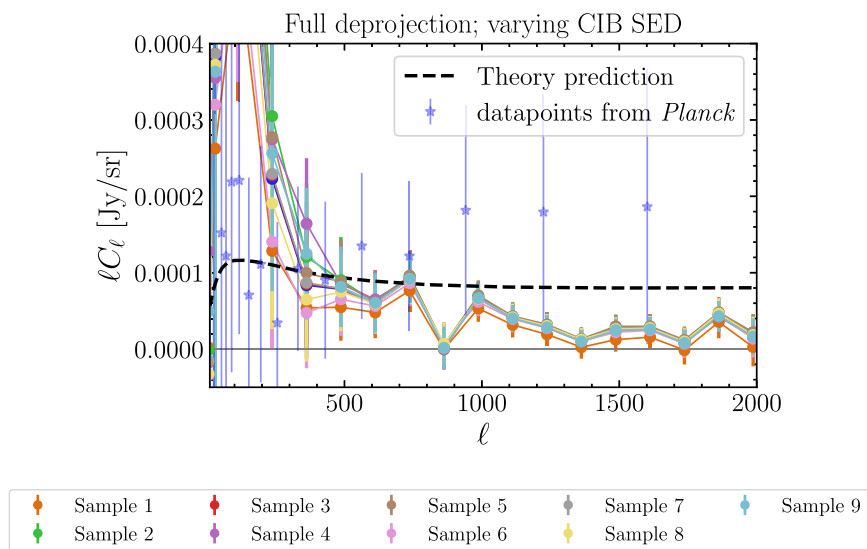

FIG. 13.   The fully moment-deprojected $(\mathrm{CIB} + \delta\beta + \delta T_{\mathrm{eff}}^{\mathrm{CIB}} + \mathrm{CMB}^5$-deprojected) y map cross-correlated with the 857 GHz CIB map from Ref. [41] for various choices of the CIB SED consistent with the posterior in Fig. 3. We also include data points from the *Planck* detection of this signal [69], digitized from Fig. 15 of that reference. The different-colored points correspond to different samples drawn from the posterior for β, T shown in Fig. 3.

we also include the data points as measured by Ref. [69] for C_ℓ^{y857}, which we have digitized from their Fig. 15.[13] As such, we hope that these maps will make possible a fruitful joint CIB-tSZ analysis.

VII. DISCUSSION

In this work, we have presented PYILC, a fully public Python package to perform NILC on full-sky HEALPix maps. We hope that PYILC will be useful both for analyses of real data to isolate maps of the CMB and tSZ signals and any other component whose spectral signature is known, as well as for analyses of *simulations*, in particular, in cases where one wants to test a signal-extraction pipeline on realistic data and to understand biases due to foregrounds. We have validated that our implementation in PYILC produces maps that closely match the official *Planck* tSZ maps (when making the same analysis choices), which were produced with a nonpublic pipeline. Our code is fully available and easily extensible to consider other sky components and other deprojection options.

We have used PYILC to perform NILC on the *Planck* NPIPE data to isolate the (almost) full-sky Compton-y signal in our Universe. We note that y maps have previously been made with this data [25,26], as well as in the official *Planck* analysis [24,37]. We have made several improvements to the analysis, including a slightly different input-map-processing step where we mask the point sources such that they do not contribute to the calculation of the NILC weights; as point sources are always masked in any final analysis mask, this does not affect any science obtained from the maps but improves their noise properties and usable sky fractions. Additionally, we provide y maps where we have explicitly removed the CIB by deprojecting both a modified blackbody component and its first moments. Our CMB-deprojected tSZ map has $\approx 10\%$ lower noise than that in the official *Planck* 2015 tSZ map.

We quantify the amount of CIB contamination in our maps by directly cross-correlating them with maps of the CIB. We find that our maps have less CIB contamination than the 2015 *Planck* NILC map. We have also presented CIB-moment-deprojected y maps, which we hope can be used to measure $\langle y - LSS \rangle$ cross-correlations without CIB bias.

We use these y maps in our companion paper [28] to detect the cross-correlation of the reconstructed CMB lensing signal and the tSZ signal, for only the second time [23], and with significantly more robustness against CIB contamination [70]. We hope that our techniques for

removing the CIB from such a measurement will be useful for a wide range of future tSZ cross-correlation analyses, including the tSZ-CIB cross-correlation [68,69].

We note that *Planck y* maps, while currently state of the art, will shortly be surpassed in signal-to-noise ratio by upcoming data from ACT [71], SPT-3G [72], and soon the Simons Observatory [73]. However, the *Planck* data will remain dominant in the Compton-y reconstruction for the foreseeable future at $\ell \lesssim 1000$, as atmospheric noise becomes too large at low multipoles for ground-based experiments to surpass *Planck* (in temperature) over this range—see, e.g., Fig. 36 of Ref. [73]. A hybrid combination of the ground-based and space-based data will thus become the standard approach in the near future (as already demonstrated in Refs. [47,74]).

VIII. SUMMARY OF PUBLIC PRODUCTS

We make various configurations of our Compton-y maps publicly available at [75]. We also release our full NILC pipeline at [76], which takes as input a set of single-frequency maps and executes the NILC algorithms described in this paper; we include as explanatory files several input files we used to construct our NILC y maps from the *Planck* data products. We note that the input is not the raw single-frequency NPIPE maps but instead the monopole- and dipole-subtracted, inpainted NPIPE maps described earlier in this work. We also release our inpainting code in the same GitHub repository as PYILC, along with our inpainting mask and several analysis masks, which are also at [77].

---

[13]Note that their points are given in units of the Compton-y parameter; we convert from Compton-y to Jy/sr using the conversion factors appropriate for 857 GHz in Table 6 of Ref. [32].

[1] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson, Astrophys. J. **142,** 414 (1965).

[2] A. A. Penzias and R. W. Wilson, Astrophys. J. **142,** 419 (1965).

[3] Y. B. Zeldovich and R. A. Sunyaev, Astrophys. Space Sci. **4,** 301 (1969).

[4] R. A. Sunyaev and Y. B. Zeldovich, Astrophys. Space Sci. **7,** 3 (1970).

[5] J. L. Puget, A. Abergel, J. P. Bernard, F. Boulanger, W. B. Burton, F. X. Desert, and D. Hartmann, Astron. Astrophys. **308,** L5 (1996).

[6] E. Dwek, R. G. Arendt, M. G. Hauser, D. Fixsen, T. Kelsall, D. Leisawitz, Y. C. Pei, E. L. Wright, J. C. Mather, S. H. Moseley *et al.*, Astrophys. J. **508,** 106 (1998).

[7] D. J. Fixsen, E. Dwek, J. C. Mather, C. L. Bennett, and R. A. Shafer, Astrophys. J. **508,** 123 (1998).

[8] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, Astrophys. J. **473,** 576 (1996).

[9] J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson, Astrophys. J. **512,** 511 (1999).

[10] D. J. Fixsen, Astrophys. J. **707,** 916 (2009).

[11] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern *et al.*, Astrophys. J. Suppl. Ser. **148,** 97 (2003).

[12] E. Martínez-González, J. M. Diego, P. Vielva, and J. Silk, Mon. Not. R. Astron. Soc. **345,** 1101 (2003).

[13] H. K. Eriksen, J. B. Jewell, C. Dickinson, A. J. Banday, K. M. Górski, and C. R. Lawrence, Astrophys. J. **676,** 10 (2008).

[14] D. S. Seljebotn, T. Bærland, H. K. Eriksen, K. A. Mardal, and I. K. Wehus, Astron. Astrophys. **627,** A98 (2019).

[15] J. Delabrouille, J. F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux, Astron. Astrophys. **493,** 835 (2009).

[16] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, Mon. Not. R. Astron. Soc. **418,** 467 (2011).

[17] J.-F. Cardoso, M. Le Jeune, J. Delabrouille, M. Betoule, and G. Patanchon, IEEE J. Sel. Top. Signal Process. **2,** 735 (2008).

[18] G. Hurier, J. F. Macías-Pérez, and S. Hildebrandt, Astron. Astrophys. **558,** A118 (2013).

[19] X. Chen and E. L. Wright, Astrophys. J. **694,** 222 (2009).

[20] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, Mon. Not. R. Astron. Soc. **410,** 2481 (2011).

[21] Y. S. Abylkairov, O. Darwish, J. C. Hill, and B. D. Sherwin, Phys. Rev. D **103,** 103510 (2021).

[22] Y. Akrami, K. J. Andersen, M. Ashdown, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak *et al.* (Planck Collaboration), Astron. Astrophys. **643,** A42 (2020).

[23] J. C. Hill and D. N. Spergel, J. Cosmol. Astropart. Phys. 02 (2014) 030.

[24] N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, N. Bartolo *et al.* (Planck Collaboration), Astron. Astrophys. **594,** A22 (2016).

[25] H. Tanimura, M. Douspis, N. Aghanim, and L. Salvati, Mon. Not. R. Astron. Soc. **509,** 300 (2022).

[26] J. Chandran, M. Remazeilles, and R. B. Barreiro, Mon. Not. R. Astron. Soc. **526,** 5682 (2023).

[27] J. Chluba, J. C. Hill, and M. H. Abitbol, Mon. Not. R. Astron. Soc. **472,** 1195 (2017).

[28] F. McCarthy and J. C. Hill, companion paper, Phys. Rev. D **109,** 023529 (2024).

[29] J. Carron, M. Mirmelstein, and A. Lewis, J. Cosmol. Astropart. Phys. 09 (2022) 039.

[30] N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro *et al.* (Planck Collaboration), Astron. Astrophys. **641,** A1 (2020).

[31] N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro *et al.* (Planck Collaboration), Astron. Astrophys. **571,** A2 (2014).

[32] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday *et al.* (Planck Collaboration), Astron. Astrophys. **571,** A9 (2014).

[33] H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje, Astrophys. J. **612,** 633 (2004).

[34] F. J. Narcowich, P. Petrushev, and J. D. Ward, SIAM J. Math. Anal. **38,** 574 (2006).

[35] M. Remazeilles, A. Rotti, and J. Chluba, Mon. Not. R. Astron. Soc. **503,** 2478 (2021).

[36] A. Kusiak, K. M. Surrao, and J. C. Hill, Phys. Rev. D **108,** 123501 (2023).

[37] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday *et al.* (Planck Collaboration), Astron. Astrophys. **571,** A21 (2014).

[38] M. P. Viero, L. Wang, M. Zemcov, G. Addison, A. Amblard, V. Arumugam, H. Aussel, M. Béthermin, J. Bock, A. Boselli *et al.*, Astrophys. J. **772,** 77 (2013).

[39] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday *et al.* (Planck Collaboration), Astron. Astrophys. **571,** A30 (2014).

[40] D. S. Y. Mak, A. Challinor, G. Efstathiou, and G. Lagache, Mon. Not. R. Astron. Soc. **466,** 286 (2017).

[41] D. Lenz, O. Doré, and G. Lagache, Astrophys. J. **883,** 75 (2019).

[42] A. Lewis and S. Bridle, Phys. Rev. D **66,** 103511 (2002).

[43] A. Lewis, Phys. Rev. D **87,** 103529 (2013).

[44] J. Torrado and A. Lewis, arXiv:1910.019.

[45] J. Torrado and A. Lewis, J. Cosmol. Astropart. Phys. 05 (2021) 057.

[46] A. Gelman and D. B. Rubin, Stat. Sci. **7,** 457 (1992).

[47] M. S. Madhavacheril, J. C. Hill, S. Næss, G. E. Addison, S. Aiola, T. Baildon, N. Battaglia, R. Bean, J. R. Bond, E. Calabrese *et al.*, Phys. Rev. D **102,** 023534 (2020).

[48] C. Shang, Z. Haiman, L. Knox, and S. P. Oh, Mon. Not. R. Astron. Soc. **421,** 2832 (2012).

[49] F. McCarthy and M. S. Madhavacheril, Phys. Rev. D **103,** 103515 (2021).

[50] R. K. Sachs and A. M. Wolfe, Astrophys. J. **147,** 73 (1967).

[51] A. Lewis, A. Challinor, and D. Hanson, J. Cosmol. Astropart. Phys. 03 (2011) 018,

[52] C. Creque-Sarbinowski, S. Bird, and M. Kamionkowski, Phys. Rev. D **94,** 063519 (2016).

[53] J. C. Hill, Phys. Rev. D **98,** 083542 (2018).

[54] D. Alonso, J. Sanchez, A. Slosar (LSST Dark Energy Science Collaboration), Mon. Not. R. Astron. Soc. **484,** 4127 (2019).

[55] G. Fabbian, J. Carron, A. Lewis, and M. Lembo, Phys. Rev. D **103,** 043535 (2021).

[56] M. Lembo, G. Fabbian, J. Carron, and A. Lewis, Phys. Rev. D **106,** 023525 (2022).

[57] K. M. Surrao, O. H. E. Philcox, and J. C. Hill, Phys. Rev. D **107,** 083521 (2023).

[58] J. C. Hill, B. D. Sherwin, K. M. Smith, G. E. Addison, N. Battaglia, E. S. Battistelli, J. R. Bond, E. Calabrese, M. J. Devlin, J. Dunkley *et al.*, arXiv:1411.8004.

[59] L. Thiele, J. C. Hill, and K. M. Smith, Phys. Rev. D **99,** 103511 (2019).

[60] B. Bolliet, B. Comis, E. Komatsu, and J. F. Macías-Pérez, Mon. Not. R. Astron. Soc. **477,** 4957 (2018).

[61] B. Bolliet, T. Brinckmann, J. Chluba, and J. Lesgourgues, Mon. Not. R. Astron. Soc. **497,** 1332 (2020).

[62] B. Bolliet, J. Colin Hill, S. Ferraro, A. Kusiak, and A. Krolewski, J. Cosmol. Astropart. Phys. 03 (2023) 039.

[63] D. Blas, J. Lesgourgues, and T. Tram, J. Cosmol. Astropart. Phys. 07 (2011) 034.

[64] N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers, Astrophys. J. **758,** 74 (2012).

[65] N. Ben Bekhti, L. Flöer, R. Keller, J. Kerp, D. Lenz, B. Winkel, J. Bailin, M. R. Calabretta, L. Dedes *et al.* (HI4PI Collaboration), Astron. Astrophys. **594,** A116 (2016).

[66] N. Battaglia, J. R. Bond, C. Pfrommer, J. L. Sievers, and D. Sijacki, Astrophys. J. **725,** 91 (2010).

[67] N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers, Astrophys. J. **758,** 75 (2012).

[68] A. Maniyar, M. Béthermin, and G. Lagache, Astron. Astrophys. **645,** A40 (2021).

[69] P. A. R. Ade, N. Aghanim, M. Arnaud, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, N. Bartolo *et al.* (Planck Collaboration), Astron. Astrophys. **594,** A23 (2016).

[70] G. Hurier, Astron. Astrophys. **575,** L11 (2015).

[71] S. W. Henderson, R. Allison, J. Austermann, T. Baildon, N. Battaglia, J. A. Beall, D. Becker, F. De Bernardis, J. R. Bond, E. Calabrese *et al.*, J. Low Temp. Phys. **184,** 772 (2016).

[72] B. A. Benson, P. A. R. Ade, Z. Ahmed, S. W. Allen, K. Arnold, J. E. Austermann, A. N. Bender, L. E. Bleem, J. E. Carlstrom, C. L. Chang *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9153,** 91531P (2014).

[73] P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso, M. A. Alvarez, K. Arnold, P. Ashton, J. Austermann *et al.*, J. Cosmol. Astropart. Phys. 02 (**2019**) 056.

[74] N. Aghanim, M. Douspis, G. Hurier, D. Crichton, J. M. Diego, M. Hasselfield, J. Macias-Perez, T. A. Marriage, E. Pointecouteau, M. Remazeilles *et al.*, Astron. Astrophys. **632,** A47 (2019).

[75] https://users.flatironinstitute.org/fmccarthy/ymaps_PR4_McCH23/.

[76] https://github.com/jcolinhill/pyilc/.

[77] https://sdsc-users.flatironinstitute.org/fmccarthy/ymaps_PR4_McCH23/.

[78] A. Zonca, L. P. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. M. Gorski, J. Open Source Software **4,** 1298 (2019).

[79] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. **622,** 759 (2005).