Reionization effective likelihood from Planck 2018 data

Chen Heinrich1,* and Wayne Hu2,3

1California Institute of Technology, Pasadena, California 91125, USA
2Kavli Institute for Cosmological Physics, Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA
3Department of Astronomy & Astrophysics, University of Chicago, Chicago, Illinois 60637, USA

(Received 11 May 2021; accepted 14 July 2021; published 2 September 2021)

We release RELIKE (reionization effective likelihood), a fast and accurate effective likelihood code based on the latest Planck 2018 data that allows one to constrain any model for reionization between 6 < z < 30 using five constraints from the CMB reionization principal components (PC). We tested the code on two example models which showed excellent agreement with sampling the exact Planck likelihoods using either a simple Gaussian PC likelihood or its full kernel density estimate. This code enables a fast and consistent means for combining Planck constraints with other reionization data sets, such as kinetic Sunyaev-Zeldovich effects, line-intensity mapping, luminosity function, star formation history, quasar spectra, etc., where the redshift dependence of the ionization history is important. Since the PC technique tests any reionization history in the given range, we also derive model-independent constraints for the total Thomson optical depth \(\tau_{PC} = 0.0619^{+0.0056}_{-0.0060}\) and its 15 \(\leq z \leq 30\) high redshift component \(\tau_{PC}(15,30) < 0.020\) (95% C.L.). The upper limits on the high-redshift optical depth is a factor of \(\sim 3\) larger than those reported in the Planck 2018 cosmological parameter paper using the FlexKnot method and we validate our results with a direct analysis of a two-step model which permits this small high-z component.

DOI: 10.1103/PhysRevD.104.063505

I. INTRODUCTION

The cosmic microwave background (CMB) has helped establish the \(\Lambda\)CDM model as the standard cosmological model [1]. While many components of this standard cosmology model are well understood, the details of the process of reionization however, remains one of the most uncertain pieces (see e.g., [2] and references therein). It is usually characterized by a single parameter, the total Thomson optical depth, whose uncertainty propagates into the inferences of other important parameters such as the primordial power spectrum amplitude. Through it, the uncertainty in reionization will become one of the major sources of uncertainty for measuring the sum of neutrino masses from future gravitational lensing measurements of the CMB [3,4] and will also have implications for inferring cosmic acceleration through the growth of structure [5].

Many probes of the reionization era exist today and with future probes it will only become more important to combine them consistently the full implications of the CMB. In the CMB data, free electrons during the reionization epoch Compton scatter with the CMB photons to cause a suppression of the CMB primary anisotropies; on the large scales, the same scattering process of the CMB temperature quadrupole in the frame of the free electrons induces an additional reionization bump in the CMB E-mode polarization. On the small scales, the kinetic Sunyaev-Zeldovich (kSZ) effects generated by the scattering of the CMB photons off ionized gas in bulk motion [6,7] are used to constrain morphology and duration of reionization [8,9]. Additionally, observations of galaxy luminosity functions, quasars, star formation rate, line-intensity mapping also provide important and complementary information on reionization [2]. In this paper, we aim to extract all the information in the large-scale E-mode observations for constraining the global ionization history of the Universe, and provide a way to consistently analyze them with all of these other probes.

In the standard CMB analysis, reionization has been modeled as a steplike transition in the global ionization history in the form of a tanh function, with the step location parametrized by the total Thomson optical depth induced. This steplike tanh model assumes, by construction, that there is negligible ionization before the transition. However, the shape of the reionization bump induced in the CMB E-mode polarization at large angles contains more information on the coarse-grained evolution of the ionization history than the total optical depth [10,11].

To extract the most information possible from this E-mode bump, Refs. [10,11] developed the principal
component (PC) method, where a few PCs is sufficient to describe the entire model space of physical ionization histories regarding their observable impact on the large-angle $E$-mode power spectrum. This method has been applied to WMAP and Planck data to obtain complete constraints on reionization models [1,11–13]. It was also adopted for a Planck 2013 analysis for marginalizing inflationary parameters in Ref. [14], for studying inflationary features vs ionization history when constraining inflationary parameters along with the kernel density estimate (KDE) technique and the Gaussian approximation [13] used for building the effective likelihoods. Then in Sec. III, we describe the PC results using the Planck 2018 + SRoll2 likelihoods, and compare with previous results in literature. In Sec. IV, we test the effective likelihood code and demonstrate its fidelity with two examples: (1) the standard steplike model; and (2) a two-parameter model where a plateau of ionization at high-$z$ is added to the standard tanh model for illustration purposes. We also derive in Sec. V model-independent results on the optical depth using the Planck analysis. Finally, we summarize our results and conclude in Sec. VI.

II. PC TECHNIQUE

The principal component technique for constraining reionization using the large-angle $C^{EE}_\ell$ polarization spectrum was first introduced in Ref. [10]. Here, we briefly summarize the PC technique itself in Sec. II A. We present techniques for building its effective likelihoods which can be used to rapidly explore any reionization model in Sec. II B. We refer the readers to Refs. [11,13] respectively for a more complete description of these topics.

A. Reionization principal components

We begin by parametrizing $x_e(z)$, the ionization fraction relative to the fully ionized hydrogen at redshift $z$, into its principal components $S_a(z)$ with respect to the CMB $E$-mode polarization:

$$x_e(z) = x_{e,\text{fid}}(z) + \sum_a m_a S_a(z),$$  \hspace{1cm} (1)

where $m_a$ are the PC amplitudes and $x_{e,\text{fid}}(z)$ is the fiducial model. We obtain the PC basis functions $S_a(z)$ as eigenfunctions of the Fisher information matrix for $x_e(z)$ in a given range $z_{\text{min}} < z < z_{\text{max}}$ from cosmic variance limited $C^{EE}_\ell$ measurements

$$F_{ij} = \sum_\ell \frac{2\ell + 1}{2} \frac{\partial \ln C^{EE}_\ell}{\partial x_e(z_i)} \frac{\partial \ln C^{EE}_\ell}{\partial x_e(z_j)} = \sum_a \frac{S_a(z_i) S_a(z_j)}{\sigma_a^2},$$  \hspace{1cm} (2)
where we have discretized the redshift space with \( \delta z = 0.25 \), and where \( \sigma_a^2 \) are the expected variances of the PCs. The Fisher matrix is computed at a fiducial model which we take to be \( x^\text{fid} (z) = 0.15 \) in the parametrized range. Note that following Ref. [18], we have updated the PCs to be computed with Planck 2015 rather than WMAP best-fit, which results in minor differences between our \( S_a \) and those of Ref. [13].

We then rank-order the PCs from low to high variance, so that for the range \( z_{\text{min}} = 6 \) (to be consistent with Ly\( \alpha \) forest constraints, e.g., [20]) and \( z_{\text{max}} = 30 \), only the first 5 components are needed to describe all the information on \( x_e \) carried by \( C_{EE}^E \) to cosmic variance limit. In the work that follows, we therefore truncate the sum at \( a = 1 \ldots N_{\text{PC}} \), where \( N_{\text{PC}} = 5 \) for \( z_{\text{max}} = 30 \) in the fiducial analysis.2

In Fig. 1 (top), we show the corresponding fiducial basis functions \( S_a(z) \). Notice that the series resembles a Fourier decomposition of \( x_e(z) \) rank ordered from low to high frequency. This is because high frequency variations in \( x_e \) leave little observable imprint in \( C_{EE}^E \). As a consequence the \( N_{\text{PC}} \) PCs are a complete representation of the observable impact of \( x_e(z) \) on the \( C_{EE}^E \), rather than the ionization history itself. In other words, given any \( x_e^{\text{true}}(z) \), we can project it onto the \( N_{\text{PC}} \) PC basis through

\[
m_a = \int_{z_{\text{min}}}^{z_{\text{max}}} dz \frac{S_a(z) [x_e^{\text{true}}(z) - x_e^{\text{fid}}(z)]}{z_{\text{max}} - z_{\text{min}}},
\]

where the reconstructed \( x_e(z) \) through Eq. (1) with truncated PCs will not reproduce the true ionization history \( x_e(z) \neq x_e^{\text{true}}(z) \), but rather it is the observed \( C_{EE}^E \) that is reproduced to cosmic variance precision. We reiterate therefore, that the PC analysis is not a tool for reconstructing the ionization history from observations, but rather a forward-modeling tool which, by reducing the dimensionality of the model space to \( N_{\text{PC}} \), allows us to constrain all possible ionization histories between \( z_{\text{min}} < z < z_{\text{max}} \) in a single analysis. For the Planck data set, most of the information on the ionization history comes from constraints on the first two modes which probe the amount of low vs high redshift optical depth. However all 5 PCs for \( z_{\text{max}} = 30 \) carry some constraint, with the higher modes controlling finer variations in redshift.

We compute the CMB power spectra with PCs using a modified version of CAMB\(^3\) [21,22] (see [13]). We follow CAMB and assume for \( z < 6 \), fully ionized hydrogen and singly ionized helium whereas for \( z \leq 3.5 \) [23], doubly ionized helium [23] with a tanh transition width \( \Delta z = 0.5 \), given the helium fraction

\[
f_{\text{He}} = \frac{n_{\text{He}}}{n_H} = \frac{m_H}{m_{\text{He}}} \frac{Y_p}{1 - Y_p},
\]

as the ratio of the helium to hydrogen number density, where \( Y_p \) is the helium mass fraction, chosen to be consistent with big bang nucleosynthesis for a given baryon density.

We obtain posterior constraints on the PC parameters \( m_a \) from Markov Chain Monte Carlo sampling using CosmoMC\(^4\) [24,25] of the relevant Planck likelihoods discussed in Sec. III in an otherwise fiducial LCDM cosmology with the standard 5 parameters: the baryon density \( \Omega_b h^2 \), cold dark matter density \( \Omega_c h^2 \), effective angular sound horizon \( \theta_{\text{MC}} \), scalar power spectrum log amplitude \( \ln(10^{10} A_s) \) and its tilt \( n_s \). We take flat uninformative priors in \( m_a \) but discuss additional priors imposed by removing demonstrably unphysical ionization histories in Sec. II C.

### B. Effective PC likelihood

The completeness property of the PCs enables us to turn PC chains obtained from a MCMC run into an effective likelihood that can be used to test any reionization model with the CMB without a cumbersome reanalysis at the level of power spectra data. In the following, we briefly recap the kernel density estimate (KDE) technique used to build this

\( ^2\)We also test the robustness of our fiducial analysis in Sec. V by taking for \( z_{\text{max}} = 50 \) where \( N_{\text{PC}} = 7 \) is required at the cosmic variance limit.

\( ^3\)CAMB: http://camb.info.

\( ^4\)CosmoMC: http://cosmologist.info/cosmomc.
likelihood, as well as a Gaussian approximation, and refer the readers to Ref. [13] for more details.

The PC chains are composed of $N_{\text{sample}}$ samples of discrete values of $\mathbf{m} = \{m_1, \ldots, m_5\}$ along with multiplicities $w_i$ for $i = 1 \ldots N_{\text{sample}}$. Given any physical ionization history $x_e(z)$, we first obtain its PC representation $\mathbf{m}$ using Eq. (3). Since $\mathbf{m}$ can take any continuous value, we approximate its effective likelihood with a kernel density estimate

$$L_{\text{KDE}}(\text{data}|\mathbf{m}) = \sum_{i=1}^{N_{\text{sample}}} w_i k_f(\mathbf{m} - \mathbf{m}_i),$$

(5)

where the overall normalization is arbitrary, and where we have chosen the smoothing kernel $k_f$ to be a multivariate Gaussian with mean zero and covariance $f \mathbf{C}$, where $\mathbf{C}$ is the $N_{\text{PC}} \times N_{\text{PC}}$ covariance matrix estimated from the PC chains (see Table I) and $f$ is a factor smaller than 1. For a Gaussian posterior, increasing the covariance by $1 + f$ corresponds to increasing the standard deviation by approximately $1 + f/2$. To minimize the amount of smoothing needed while still maintaining good accuracy in the tail of the distribution for any physical models, we oversample the PC chains by running the PC chains far beyond convergence for about $N_{\text{sample}} \approx 10^6$ chain samples. For this $N_{\text{sample}}$, a fraction of $f = 0.14$ should be sufficient for most models, and indeed we have tested it to work well for the ones we study in Sec. IV.

An even simpler effective likelihood is the Gaussian approximation of the $m_a$ posterior using the PC mean $\bar{\mathbf{m}}$ and covariance $\mathbf{C}$

$$L_{\text{Gauss}}(\text{data}|\mathbf{m}) = \frac{e^{-\frac{1}{2}(\mathbf{m} - \bar{\mathbf{m}})^T \mathbf{C}^{-1} (\mathbf{m} - \bar{\mathbf{m}})}}{(2\pi)^{N_{\text{PC}}} |\mathbf{C}|}.$$  

(6)

For any set of model parameters $\mathbf{p}$ with their given prior probability distribution $P(\mathbf{p})$, their posterior distribution can be approximated as

$$P(\mathbf{p}|\text{data}) \propto L_{\text{PC}}[\text{data}|\mathbf{m}(\mathbf{p})]P(\mathbf{p}).$$

(7)

with PC = KDE or Gaussian. This approach obviates the need to implement specific reionization models into CAMB and conduct separate CosmoMC sampling for each, thereby significantly reducing both human and computational effort. In Sec. IV we test the effective KDE and Gaussian likelihoods against an exact implementation for two example models: the standard steplike model and a two-step model allowing for high-redshift ionization.

C. Cumulative optical depth

Although reionization PCs are mainly a tool for model testing using complete information from large angle CMB polarization, as they do not reconstruct the rapidly varying aspects of $x_e(z)$ itself, they do provide model-independent constraints on redshift integrated quantities such as the cumulative Thomson optical depth

$$\tau(z, z_{\text{max}}) = n_H(0)\sigma_T \int_z^{z_{\text{max}}} dz x_e(z)(1 + z)^2 \frac{H(z)}{\sigma},$$

(8)

where $n_H(0)$ is the hydrogen number density at $z = 0$, $\sigma_T$ is the Thomson scattering cross section and $H(z)$ is the Hubble parameter. Given the tight constraints on cosmological parameters in the $\Lambda$CDM model, this quantity is well approximated by

$$\tau_{\text{PC}}(z, z_{\text{max}}) = \sum_{a=1}^{N_{\text{PC}}} m_a \tau_a(z, z_{\text{max}}) + \tau_{\text{fid}}(z, z_{\text{max}}),$$

(9)

where $\tau_a$ and $\tau_{\text{fid}}$ are defined by employing Eq. (1) in Eq. (8) using the fiducial model for $H(z)$ and number densities. This differs from Ref. [13] where their cosmological dependence was retained in determining $\tau_{\text{PC}}$. When we omit the redshift arguments, we implicitly mean the total range, e.g., $\tau_{\text{PC}} \equiv \tau_{\text{PC}}(0, z_{\text{max}})$.

In Fig. 1 (bottom), we display $\tau_{\text{PC}}(z, z_{\text{max}})$ for the $z_{\text{max}} = 30$ PCs. Notice that positive values of the first component $m_1$ mainly represents the optical depth that is accumulated at high redshift whereas negative values of $m_2$ provides the same for low redshift. Since these two are the best measured components, the CMB mainly determines the amount of low vs high redshift ionization. The underlying reason is that in $C_l^{EE}$ the higher the redshift, the larger the relative contribution at higher $l$ due to the size of the horizon when the photons scattered. For the less well measured PCs with $a > 2$, the contributions to the total optical depth rapidly diminish and represent finer distinctions in redshift for the cumulative optical depth and multipole for the underlying $C_l^{EE}$ spectrum.

In the usual approach to constraining reionization in $\Lambda$CDM, one places an implicit prior on the shape of the cumulative optical depth by choosing a “tanh” or near step

| $m_1$  | $m_2$  | $m_3$  | $m_4$  | $m_5$  |
|--------|--------|--------|--------|--------|
| 1.000  | 0.657  | −0.208 | −0.094 | 0.067  |
| 0.657  | 1.000  | 0.048  | −0.273 | 0.142  |
| −0.208 | 0.048  | 1.000  | 0.064  | −0.018 |
| −0.094 | −0.273 | 0.064  | 1.000  | −0.195 |
| 0.067  | 0.142  | −0.018 | −0.195 | 1.000  |

| $\bar{m}_a$ | $\sigma(m_a)$ | $\tau_a$ | $\tau_{\text{fid}}$ |
|-------------|--------------|--------|-----------------|
| −0.116      | 0.016        | 0.078  | −0.077          |
| 0.030        | 0.063        | 0.098  | 0.131           |
| 0.29403     | −0.11227     | 0.04506| −0.01898        |
| 0.00960     |              |        |                 |
function reionization (see Sec. IVA) and then extracts a single constraint on the total optical depth. The PCs, on the other hand, allow for arbitrary values of $x_e(z)$ when no range-bound prior constraints on the mode amplitudes $m_a$ are imposed. The one subtlety when placing constraints on the cumulative optical depth is that the analysis also allows unphysical ionization fractions where $0 \leq x_e \leq x_e^{\max}$ is not satisfied. Therefore, when determining cumulative optical depth constraints as opposed to using reionization PCs as a tool for testing models, we do impose a prior by truncating the posteriors of $m_a$ after obtaining them from Markov Chain Monte Carlo sampling, following Ref. [12]

$$\sum_{a=1}^{N_{\text{PC}}} m_a^2 \leq (x_e^{\max} - x_e^{\text{fid}})^2,$$

and $m_a^- \leq m_a \leq m_a^+$ with

$$m_a^\pm = \int_{z_\text{min}}^{z_\text{max}} dz [S_a(z)][x_e^{\max} - 2x_e^{\text{fid}}(z)] \pm x_e^{\max} |S_a(z)|/2(z_\text{max} - z_\text{min}).$$

Here we take $x_e^{\max} = 1 + \bar{f}_\text{He}$ to account for the contribution of singly ionized helium. In the fiducial analysis, the prior on the sum of squares in Eq. (10) is strictly weaker than the individual priors in Eq. (11). Henceforth we refer to Eq. (11) as the physicality prior. We study its impact and robustness in Sec. V. Note that the original $0 \leq x_e \leq x_e^{\max}$ condition cannot be strictly enforced here because we keep only the first $N_{\text{PC}}$ PCs, so these priors in $m_a$ are necessary but not sufficient conditions for a physical ionization history. In other words, no physical model is excluded by these priors but some unphysical models are included. In Sec. VA we further explore the role of these and other choices of PC priors, and in Sec. IV we test PC constraints against exact constraints for two example models.

III. PLANCK 2018 PC RESULTS

We now present the complete reionization constraints obtained from the Planck 2018 likelihoods using the principal components.

A. Constraints on principal components

We use the official Planck likelihoods [26] plik_lite_TTTEEE for the high-$\ell$ TT, TE and EE as well as lowl for the low-$\ell$ TT throughout this paper. We have tested that our results do not change if in lieu of plik_lite_TTTEEE we used the full plik_full_TTTEEE likelihood in which the foreground parameters have not been marginalized over. For the low-$\ell$ EE likelihood, we use the third-party released SRoll2 likelihood [19] in our official PC results and the effective likelihood code. In comparison with the official Planck-released SimAll likelihood, the SRoll2 likelihood has improved foreground cleaning, which enables stronger reionization constraints, especially from the low multipole moments that are important for the low redshift constraints.

In Fig. 2, we show the 1D posterior and 2D 68% and 95% confidence level contours for the 5 PC amplitudes that describe ionization models up to $z_{\text{max}} = 30$. We also show the trajectory of the tanh model (see Sec. IVA) through this space as well as the best fit tanh model points. The gray shaded region represents the parameter space that would be excluded by an additional physicality prior from Eq. (11), which is just beyond the border of the displayed regime for all but $m_1$. Results using the Planck 2018 data are shown in blue.

While all five PC amplitudes are constrained by Planck, the first two are particularly well constrained. Contrary to the Planck 2015 results of Ref. [13] shown in red, the 2018 data prefers a much smaller $m_1$, leading to a reduced optical depth, especially at high redshift. This preference also increases the impact of placing a physicality prior at low values of $m_1$ when deriving model-independent constraints as we shall discuss further in Sec. V. We show a zoomed version of the $m_1 - m_2$ plane in Fig. 3. In the physicality prior excluded region (gray), the ionization and the cumulative optical depth at high redshift becomes negative.

In Fig. 3, we also show the Planck 2018 constraints using the older SimAll likelihood. The improvement from switching to the SRoll2 likelihood mainly strengthens the upper bound in the best constrained direction in the $m_1 - m_2$ plane which shifts constraints toward higher total optical depth (see also Fig. 11).

In Fig. 2, we also show the constraints assuming a Gaussian posterior in $m_a$ using the PC chain covariance and means listed in Table I. The visual agreement with the full posteriors is remarkably good, especially in the well-constrained components. This suggests that the simple Gaussian likelihood of Eq. (6) suffices for most constraints on models. We shall see in Sec. IV that a quantitative comparison indeed shows good agreement between the KDE and Gaussian likelihoods, with only minor differences in the shape of their posterior inferences.

IV. MODEL TESTING WITH PCS

We now turn to constraining reionization models with PCs using the effective likelihood approach described in §IV. The benefit of this approach is that complete constraints on any model of the ionization history between $0 \leq z \leq z_{\text{max}}$ are obtained ahead of time and compressed into a simpler likelihood, so that when doing model testing

---

5The Planck 2015 best-fit for the tanh model is derived from a chain best-fit model as taken from Ref. [13], whereas for the Planck 2018 point in this paper we used a minimizer to find the best-fit; there are also slight model differences, where in Planck 2018 we used a tanh width of $\Delta z = 0.015(1 + z)$ instead of $\Delta z = 0.5$ for the Planck 2015 best-fit. Using $\Delta z = 0.5$ for Planck 2018 moves the cross negligibly on this plot.
one no longer needs to modify the $x_e(z)$ function and jointly sample the other cosmological parameters in for example CAMB, each time a new ionization model needs to be tested. These provide savings on human time in addition to the computing speed-ups from the likelihood evaluations.

Our code is available on GitHub at https://github.com/chenheinrich/RELIKE. The initial release corresponds to the fiducial results in this paper, i.e., derived from the Planck 2018 plik_lite_TTTEEE+lowl+SRoll2 likelihoods at $z_{\text{max}} = 30$. This is also the default analysis for all results that follow except where noted. The most important part of the code release is the python package called RELIKE which provides the PC parameters for a model and also implements the Gaussian effective likelihood. The user can specify a customized $x_e(z)$ function taking in reionization model parameters and use RELIKE to return the effective likelihood of the model. The returned likelihood is already marginalized over cosmological parameters. To obtain marginalized distributions more easily for high-dimensional parameter space, one may choose to sample RELIKE within a MCMC sampler such as Cobaya\textsuperscript{6} [27] or CosmoSIS\textsuperscript{7} [28].

---

\textsuperscript{6}Cobaya: https://cobaya.readthedocs.io/en/latest/.

\textsuperscript{7}CosmoSIS: https://bitbucket.org/joezuntz/cosmosis/wiki/Home.
We also provide a pre-assembled CosmoMC code using the CosmoMC’s generic sampler, with which we have generated the results presented in this section. This CosmoMC code has an internal implementation of both the Gaussian and KDE likelihoods in fortran, which is separate from our main python package that only supports the Gaussian likelihood.

When using the KDE likelihood, we suggest using the default value of \( f = 0.14 \) to avoid over-smoothing parameter posteriors while maintaining accuracy during the KDE operation. All our KDE results in this section are computed with \( f = 0.14 \), which we shall soon show to work well in recovering results from an exact MCMC. Note that in both the Gaussian and the KDE modes, the model parameters or priors must be arranged to explicitly satisfy fully-ionized hydrogen and singly ionized helium for \( z \leq 6 \).

Next we demonstrate the use and the successful recovery of parameter posteriors using the effective likelihood code by comparing them to an exact MCMC analysis using the original Planck likelihoods for two examples: (1) The tanh model, which is the standard approach adopted in CAMB; (2) a two-step toy model which has one additional parameter allowing for a high-\( z \) ionization plateau before the transition to full ionization in the canonical tanh model.

### A. Example 1: tanh model

The steplike model used as the standard approach in CAMB describes the hydrogen and singly ionized helium reionization as a tanh function:

\[
x_e(z) = \frac{1 + f_{\text{He}}}{2} \left\{ 1 + \tanh \left[ \frac{y(z_e) - y(z)}{\Delta y} \right] \right\} + x_e^{\text{rec}},
\]

with \( y(z) = (1 + z)^{3/2} \), \( \Delta y = (3/2)(1 + z)^{1/2}\Delta z \). Instead of using \( \Delta z = 0.5 \) as in the standard setting by CAMB, we adopt a smaller width \( \Delta z = 0.015(1 + z) \) so as to better capture the required full ionization below \( z = 6 \). Here \( x_e^{\text{rec}} \) is the ionization history from recombination only and we follow the fiducial prescription described in Sec. II for the doubly ionized helium transition at \( z = 3.5 \).

The tanh model is parametrized by its total optical depth for which we take a flat prior with an additional constraint that \( z_e \geq 6.1 \) so as to satisfy the required full ionization by \( z_{\text{min}} = 6 \) given the tanh width to good approximation.

In Fig. 2, we show the trajectory that the tanh model takes in the PC space. Notice that this trajectory passes near the center of all constraints for Planck 2018 but only through the tails for some of the Planck 2015 constraints. Correspondingly, whereas the tanh model was marginally disfavored with Planck 2015 data, that is not the case with Planck 2018 data.

In Fig. 4 we show the posterior distribution of \( \tau \) for the effective likelihood results, in which the only sampled parameter is \( \tau \) (all other cosmological parameters are fixed at the Planck best-fit model), vs the exact Planck likelihood results in which the five other \( \Lambda \)CDM parameters were also varied using their standard priors. More specifically, we find

\[
\begin{align*}
\tau_{\text{tanh}} &= 0.0591^{+0.0054}_{-0.0068}, \quad \text{(Exact)} \\
\tau_{\text{tanh}} &= 0.0591^{+0.0061}_{-0.0066}, \quad \text{(KDE)} \\
\tau_{\text{tanh}} &= 0.0592^{+0.0062}_{-0.0062}, \quad \text{(Gaussian)},
\end{align*}
\]

where our exact result is also consistent with the tanh model constraint in Ref. [29] for the same combination of likelihoods.
The posteriors of all three likelihoods agree well. The KDE likelihood gives a slightly broader posterior because the kernel density estimate in the PC space introduces smoothing. We expect the smoothing to be roughly at the $\sim 7\%$ level in the parameter posterior corresponding to the smoothing factor of $1 + f = 1.14$ we applied to the PC covariance during KDE. Indeed, the KDE result has a 4% larger 68% confidence region than the exact likelihood result.

Instead of the KDE likelihood, we can alternatively use the Gaussian approximation given in Eq. (6). The Gaussian likelihood is faster to evaluate as it does not require looping over the entire $\sim 10^6$ points in the PC chain. The Gaussian likelihood result above gives symmetric tanh $\tau$ constraints where as the KDE result is able to capture some of the skewness in the exact distribution. The shift in the peaks seen in Fig. 4 between Gaussian and the exact likelihoods, $\Delta \tau \sim 0.001$, is consistent with the level of asymmetric error bars in the latter.

We also show in Fig. 5 the cumulative optical depth obtained from directly integrating the tanh model in the exact likelihood chains vs that from the 5-PC representation of the tanh models in the effective likelihood chains using Eq. (9). Notice that because of the truncation of the PC expansion at $N_{\text{PC}} = 5$, this representation shows a small oscillatory cumulative $\tau$ at $z \gtrsim 10$ whereas the true model has negligible ionization there. This is because the best constrained modes only represent the smooth features in the ionization history and truncation is similar to a low pass filter in Fourier space. Though these zero mean, small amplitude oscillations do not bias results when averaged over redshift, this effect should be born in mind when interpreting the model-independent results in Sec. V.

To quantify these residual oscillations we consider the 95% upper limit on the high-redshift optical depth. For both effective likelihoods $\tau_{\text{tanh}}(15, 30) < 0.002$ while $\tau_{\text{tanh}}(15, 30) \approx 0$ from the exact likelihood. By construction, the tanh model does not allow for a significant fraction of the total optical depth to come from high redshift. As we shall see in the next example the data themselves do allow for a small but significantly larger fraction of the total optical depth to originate from high redshift, when the form of the model permits it.

B. Example 2: The two-step model

The second example is a two-step model, where an additional step is used to give a high-$z$ ionization plateau:

$$x_e(z) = \frac{1 + f_{\text{HeHe}} - x_{e,\text{min}}}{2} \left( 1 + \tanh \left( \frac{y(z_{\text{rec}}) - y(z)}{\Delta y} \right) \right) + \frac{x_{e,\text{min}} - x_{e,\text{rec}}}{2} \left( 1 + \tanh \left( \frac{z_{\text{hi}} - z}{\Delta z_{\text{hi}}} \right) \right) + x_{e,\text{rec}},$$

(14)

where $y(z) = (1 + z)^{3/2}$, $\Delta y = (3/2)(1 + z_{\text{rec}})^{1/2} \Delta z_{\text{rec}}$ where $\Delta z_{\text{rec}} = 0.015(1 + z_{\text{rec}})$ just as in the first example for a sharper step. We choose the second step at $z_{\text{hi}} = 28$ with $\Delta z_{\text{hi}} = 1.0$, since the effective likelihood is limited to models with ionization below $z_{\text{max}} = 30$. In comparison to tanh, a single parameter $x_{e,\text{min}}$ is added in this model to control the high-$z$ ionization plateau for $z_{\text{rec}} \lesssim z \lesssim z_{\text{hi}}$. We would recover the standard tanh as $x_{e,\text{min}}$ approaches the negligible recombination value $x_{e,\text{rec}}$. We show examples of this two-step model in Fig. 6.

![Figure 5](image5.png)

**FIG. 5.** Cumulative optical depth in the tanh model as in Fig. 4 (68% and 95% C.L. contours). Overall agreement is good in regions where the contribution to $\tau_{\text{tanh}}$ is high but the model form forbids high-$z$ contributions. Oscillations around zero at high-$z$ in the KDE and Gaussian PC representations are due to the effective low pass filter from the truncation at 5 PCs which does not produce any observable difference in the CMB.

![Figure 6](image6.png)

**FIG. 6.** Two-step ionization history $x_e(z)$ for the model of Sec. IV B. In addition to a tanh transition at low-redshift as in Sec. IV A, there is an additional ionization plateau at high redshifts with a fixed transition at $z_{\text{hi}} = 28$ and $\Delta z_{\text{hi}} = 1$. Here we show the best-fit model in the Planck 2018 data, $\tau_{\text{hi}} = (0.053, 0.006)$, as well as a model on the 95% C.L. contour of Fig. 7 which maximizes $\tau_{\text{hi}} = (0.043, 0.026)$. 

063505-8
In practice, we parameterize the model by $\tau_{lo}$ and $\tau_{hi}$. They are defined as in Eq. (8) but with the boundaries $(0, z_{split})$ and $(z_{split}, \infty)$ for $\tau_{lo}$ and $\tau_{hi}$ respectively, where the split is chosen as $z_{split} = z_{re} + \Delta z_{re}$ to be conservative on reducing the preference of $\tau_{hi}$ in the data. Given a model $(\tau_{lo}, \tau_{hi})$, we find the corresponding $(z_{re}, x_e^{\text{min}})$ through an iterative search similar to that used in the canonical tanh model for finding $z_{re}$ given a total optical depth $\tau$. For both the full likelihood and effective likelihood runs, we adopt the flat priors on $\tau_{lo}$, $\tau_{hi} \in [0, 0.35]$. We further apply a prior cut in $z_{re}$ space to keep models with $z_{re} > 6.1$ for both the full and effective likelihood chains as in the tanh model to ensure full reionization by $z = 6$ to good approximation.

In Fig. 7, we display the posteriors of $\tau_{lo}$ and $\tau_{hi}$. The exact, KDE and Gaussian PC treatments again agree very well. Note that in the canonical tanh model $\tau_{lo} \lesssim 0.04$ roughly corresponds to $z_{re} < 6$ so that the posteriors are effectively cut off near this value. The models in Fig. 6 are represented by the star and cross symbols for the best fit case and a model allowed at the 95% C.L. respectively.

The marginalized 1D constraints from the exact, KDE and Gaussian likelihoods also agree well:

$$
\begin{align*}
\tau_{lo} &= 0.0526^{+0.0058}_{-0.0058}, \quad \tau_{hi} < 0.021 \ (95\% \text{ C.L.}), \quad \text{(Exact)} \\
\tau_{lo} &= 0.0525^{+0.0062}_{-0.0079}, \quad \tau_{hi} < 0.021 \ (95\% \text{ C.L.}), \quad \text{(KDE)} \\
\tau_{lo} &= 0.0528^{+0.0064}_{-0.0077}, \quad \tau_{hi} < 0.021 \ (95\% \text{ C.L.}), \quad \text{(Gaussian)}.
\end{align*}
$$

(15)

Note that the upper limit on $\tau_{hi}$ presented here is the one-sided 95% C.L. limit in 1D; whereas the model shown in Figs. 6 and 7 is located on the 95% C.L. contour in 2D and has therefore a larger $\tau_{hi} = 0.026$.

Moreover, we can derive the constraints on the total optical depth, which are also consistent between the three likelihoods as well:

$$
\begin{align*}
\tau_{2\text{step}} &= 0.0621^{+0.0062}_{-0.0061} \quad \text{(Exact)} \\
\tau_{2\text{step}} &= 0.0619^{+0.0063}_{-0.0063} \quad \text{(KDE)} \\
\tau_{2\text{step}} &= 0.0618^{+0.0062}_{-0.0062} \quad \text{(Gaussian)}.
\end{align*}
$$

(16)

Given the small amount of $\tau_{hi}$ still allowed, the total optical depth constraints in the two-step model are correspondingly slightly larger than $\tau_{\text{amb}}$. This is consistent with other examples in the literature where extended reionization to high redshifts were allowed (e.g., Refs. [30,31]). More generally, models which are sufficiently flexible at allowing for high-redshift ionization will have a comparable difference in their total optical depth.

Finally the 95% C.L. upper limit on high-redshift optical depth are also consistent:

$$
\begin{align*}
\tau_{2\text{step}}(15, 30) &< 0.014 \ (95\% \text{ C.L.}), \quad \text{(Exact)} \\
\tau_{2\text{step}}(15, 30) &< 0.016 \ (95\% \text{ C.L.}), \quad \text{(KDE)} \\
\tau_{2\text{step}}(15, 30) &< 0.016 \ (95\% \text{ C.L.}), \quad \text{(Gaussian)}.
\end{align*}
$$

(17)

The small difference in PC results vs exact reflects the mild oscillatory PC artifact around $z \sim 15$ in Fig. 8, where we show the cumulative optical depth constraints on this two-step model in the exact vs the PC likelihood construction. Note that the 95% upper limit on $\tau(15, 30)$ reported in the text are one-sided upper limits, whereas the contours in the cumulative optical depth figures enclose their respective C.L. Note also that the 95% C.L. allowed two-step model in Fig. 6 has $\tau_{2\text{step}}(15, 30) = 0.017$.

The Planck 2018 cosmological parameter paper [1] also constrained $\tau(15, 30)$ in a model-independent way using the FlexKnot method, in which $x_e(z_i)$ at the $i$th knots are varied as free parameters as well as the number of total knots. They obtained $\tau(15, 30) < 0.007 \ (95\% \text{ C.L.})$, which would strongly rule out models that are still allowed at the 95% C.L. in this class. In principle, the FlexKnot method allows for all possible physical ionization histories that increase monotonically with redshifts, and would include the two-parameter model described here. But apparently the implicit prior employed by the method disfavors optical depth at high redshift more than a flat prior in $\tau_{hi}$ does. We shall see in §V that our model independent PC approach
models such a flat prior better and of course the effective likelihood approach allows for any desired prior.

V. MODEL-INDEPENDENT RESULTS FOR OPTICAL DEPTH

While the PC approach to constraining reionization should mainly be used for testing wide classes of models efficiently and with their own physically motivated priors, it can also be employed to extract model-independent constraints on the optical depth. In this section we present these model-independent results, first on the total optical depth and then on the cumulative optical depth at high redshift, and discuss their robustness to priors on physicality of the ionization history and redshift range of the fiducial analysis.

A. Total optical depth

Using Eq. (9), we can derive constraints on the total optical depth \( \tau_{PC} = \tau_{PC}(0, z_{\text{max}}) \) with the additional physicality priors on the PC amplitudes \( m_a \) from Eq. (11). These correspond to

\[
\tau_{PC} = 0.0619^{+0.0056}_{-0.0068}, \quad (18)
\]

and the full posterior is shown in Fig. 9. Notice that the distribution matches that of the two-step model from Sec. IV B but almost 0.5\( \sigma \) shifted higher than that of the tanh model from Sec. IV A.

The similarity and difference between these results come down to implicit and explicit priors. As we have seen, results based on the tanh model requires that reionization happens suddenly at \( z_{re} \) and is maintained thereafter. Therefore, the model has a strong implicit prior that a given a total optical depth \( \tau \) comes from \( z \lesssim z_{re} \) regardless of whether the data allow or prefer ionization at higher \( z \). While this prior may be well motivated theoretically, it should be distinguished from constraints that are data driven. This was especially apparent with the Planck 2015 data where the PC constraints gave \( \tau_{PC} = 0.092 \pm 0.015 \) which differed substantially from the tanh \( \tau_{PC} = 0.079 \pm 0.017 \) [13]. As we shall discuss below, these differences were even more striking in the implications for the cumulative \( \tau \) at high redshift as we have also concretely illustrated in the two-step model of Sec. IV B.

PC constraints on the total optical depth do not suffer from assuming a specific class of ionization histories as they parameterize any \( x_e(z) \) from \( z_{\text{min}} = 6 \) to \( z_{\text{max}} = 30 \). They do of course come with their own priors namely the flat range-bound ones from Eq. (11). These have in fact caused some confusion in the literature when interpreted in terms of the total optical depth [17]. Here we first review the impact of flatness in PC space as discussed extensively in Ref. [18] and then highlight the more important role of physicality priors with Planck 2018 data as compared with 2015 data.

We first note that by virtue of Eq. (9), the total optical depth receives contribution from all PCs. However the PCs are rank ordered by their expected constraining power and so most of the information on the total optical depth comes from the first two components

\[
\tau_{PC} \approx \tau_{12} \equiv m_1 \tau_1 + m_2 \tau_2 + \left( \sum_{a=3}^5 \bar{m}_a \tau_a + \tau_{\text{fid}} \right), \quad (19)
\]
where the term in parentheses is a fixed quantity (see Table I) whose purpose is to remove offsets due to the arbitrary choice of the fiducial ionization model. We explicitly demonstrate this approximate equivalence in Fig. 10 by comparing the posterior distributions of \( \tau_{PC} \) and \( \tau_{12} \) both with and without the physicality prior.

Due to this linearity, flat priors in \( m_1 \) and \( m_2 \) correspond to flat priors in \( \tau_{12} \) for any linear trajectory through the space. For example, the tanh model trajectory is shown in Fig. 2 and is nearly linear through the Planck 2018 allowed region in the \( m_1 - m_2 \) plane. Even without the explicit construction of a flat tanh \( \tau \) prior in Sec. IVA, the implicit prior from \( m_1, m_2 \) is already nearly flat.

Reference [17] raised a potential objection regarding more general models. As illustrated in Fig. 11, lines of constant \( \tau_{12} \) are rotated compared with the \( m_1, m_2 \) axis. Therefore, the range-bound flat prior in \( m_1, m_2 \) (blue square) is not flat in its marginal \( \tau_{12} \) distribution due to the extent of the allowed parameter space along lines of constant \( \tau_{12} \): At the implied lowest \( \tau_{12} = 0.023 \) of the range in Fig. 11, this extent vanishes and then increases linearly until \( \tau_{12} = 0.130 \). Reference [17] suggests inverting this prior point-by-point in the \( m_1, m_2 \) parameter space by dividing by the implied prior \( P(\tau_{PC}) \approx P(\tau_{12})/C_{138} \); this reweighting was implemented in the Planck 2018 paper on cosmological parameters [1]. First, notice that this inversion would produce a non-flat prior on \( \tau \) for the tanh model that disfavors high values of \( \tau \). This is because the additional parameter space allowed by the range-bound \( m_n \) prior is excluded by the functional form of the model, even though the reweighting occurs globally and includes the tanh trajectory as well.

More generally, to the extent that the data constrain \( m_1 \) and \( m_2 \) better than the flat range-bound priors, these priors become irrelevant and the reweighting of Ref. [17] becomes erroneously motivated. To see this, consider a range-bound prior that is rotated to be oriented along constant \( \tau_{12} \) (see Fig. 11, blue dashed rectangle). This rotated prior in \( m_1 - m_2 \) space corresponds to a constant \( P(\tau_{12}) \) once the direction orthogonal to \( \tau_{12} \) is marginalized. Had the data constraint been shifted to slightly higher \( m_1 \) so that both prior ranges encompassed all of the 95% C.L. allowed region they both would yield the same constraints.
at that level. However the reweighting by $1/P(\tau_{12})$ would change the original prior but not the rotated one and destroy this desired equivalence. More generally the point is that for a multidimensional space, constructing a flat marginal 1D distribution per parameter in the absence of data does not guarantee a locally flat prior in the region supported by the data (see [18] for further discussion).

The real impact of our range-bound $m_1$, $m_2$ priors comes from the fact that with the Planck 2018 data, the physicality prior on $m_1$ removes nearly half of the space that would otherwise be allowed by the data at 95% C.L. This is in contrast to the Planck 2015 data where only a very small region is removed (see Fig. 3). This is reflected in the upward shift of the $\tau_{12}$ posterior in Fig. 10 with and without the physicality prior on $m_1$. This prior excludes very negative $m_1$ and we can see from Fig. 1 that these cases would have an unphysically negative ionization and cumulative optical depth at high $z$. More precisely, even though a very positive value of $m_2$ can restore physicality at the highest redshifts, the fact that there is a node near $\tau_2(15, z_{\text{max}})$ means that no choice can counteract $m_1$ there. Correspondingly, the impact of the $m_1$ prior is to remove cases where the high redshift contribution is negative which would otherwise broaden and shift the $\tau_{12}$ posterior to lower values.

This also shows how our range-bound priors are conservative: all the cases that are excluded are definitely unphysical but some of the cases that are included may also be unphysical since their physicality depends on higher-order PC components. The impact on optical depth however is small and at most, contributes a small bias that conservatively lowers the high redshift contribution which is already a small contribution to the total. This can be seen in the comparison with the two-step model in Fig. 9 and as we shall return to this point in the cumulative optical depth constraints in the next section where the high and low redshift contributions are manifest. In any case, using PCs to test explicitly physical models as in Secs. IV A and IV B cannot be affected by our conservative physicality cut.

Finally, we check that our results [Eq. (18)] are robust to extending the PC range to $z_{\text{max}} = 50$ which is negligibly different

$$
\tau_{PC} = 0.0626^{+0.0061}_{-0.0072} \quad (z_{\text{max}} = 50).
$$

Reverting the Planck 2018 likelihood to SimAll at $z_{\text{max}} = 30$ would give

$$
\tau_{PC} = 0.0582^{+0.0072}_{-0.0083} \quad \text{(SimAll)}.
$$

As with the tanh model, the SRoll2 likelihood provides tighter constraints that shift the distribution toward higher optical depth.

![Diagram](image-url)

**FIG. 12.** Cumulative optical depth $\tau_{PC}(\tilde{z}, \tilde{z}_{\text{max}})$ from the model-independent PC approach with and without the physicality prior. The prior helps exclude models with negative contributions at high-$\tilde{z}$ but conservatively allows some negative values at $\tilde{z} \sim \tilde{z}_{\text{max}} = 30$ where the oscillations in the PCs no longer cancel due to the cumulative integration. Upper limits are nearly unaffected.

### B. Cumulative optical depth

The value of a model-independent approach is even more apparent for constraints on the cumulative optical depth $\tau(\tilde{z}, \tilde{z}_{\text{max}})$ at high $\tilde{z}$. The main results are again related to the $m_1$ and $m_2$ PC constraints, both of which are well constrained. Consequently, as shown in Fig. 11, beyond the well-constrained total optical depth [as represented by the $\tau_{12}$ linear combination in Eq. (19)] is a somewhat weaker constraint in the constant $\tau_{12}$ direction that reflects whether that $\tau_{12}$ comes mainly from high or low redshift (see Sec. II C). Any finer redshift detail in the ionization history is much more poorly constrained.

In Fig. 12, we show the 68% and 95% confidence level contours for $\tau_{PC}(\tilde{z}, \tilde{z}_{\text{max}})$ with and without the physicality prior on $m_1$. Notice that without the prior, a larger range of unphysically negative optical depth is allowed, whereas with the prior it is mostly eliminated except for the highest redshifts. Moreover, upper bounds on the optical depth at high redshift are largely unchanged. These upper limits allow some contribution at high-$\tilde{z}$ but do not require it. For example,

$$
\tau_{PC}(15, 30) < 0.020 \quad \text{(95% C.L.)}.
$$

Recall again that we are reporting the one-sided upper limit here in the text, which is slightly lower than values at the contours shown in the figures. This should be contrasted with the tanh model where high-$\tilde{z}$ ionization is forbidden by the functional form of the ionization history. On the other hand the bound is comparable to that obtained for the two-step model [see Eq. (17)]. It is actually slightly larger given that the form of the two-step model still requires a steadily
decreasing cumulative optical depth at $z > z_{\text{re}}$ with the additional allowed contributions reflected in $\tau_{\text{hi}}$.

These results are in contrast to those of the Planck 2015 analysis. In Fig. 13 top panel, we compare the two. Whereas the Planck 2015 has only upper limits on the high redshift optical depth, Planck 2015 actually preferred finite optical depth at $z \gtrsim 15$.

In the lower panel, we compare the Planck 2018 results for the two different low-$\ell$ EE likelihoods. Improvements from SRoll2 come mainly from improving the lower bound at low redshifts, consistent with the improved constraints at the lowest multipoles of $C_l^{EE}$.

We also show using a separate PC chains with $z_{\text{max}} = 50$ that there is no evidence for optical depth at $z > 30$. In Fig. 14, we show that the 68% and 95% C.L. contours of $\tau(z, z_{\text{max}})$ are consistent between the $z_{\text{max}} = 30$ (red regions) and $z_{\text{max}} = 50$ results (black lines). For the cumulative high-redshift optical depth,

$$\tau_{\text{PC}}(15, 50) < 0.019 \ (95\% \ C.L.),$$

we also obtain consistent upper limits to the $z_{\text{max}} = 30$ case of Eq. (22).

## VI. CONCLUSION

In conclusion, we have used constraints on reionization principal components of the large scale $EE$ polarization to produce a fast and accurate effective likelihood for the final release of Planck 2018 data. We replaced the official SimAll likelihood for low-$\ell$ EE with the SRoll2 likelihood independently released in 2019 which uses improved foreground cleaning methods. Our effective likelihood code package is called RELIKE and is available publicly on Github. We expect the code to facilitate the testing of any global ionization history between $6 < z < 30$ with the Planck data and to enable fast and consistent joint analyses with other reionization datasets.

To test this effective likelihood code, we used two examples: 1) a canonical tanh model with a single transition redshift; and 2) a two-parameter toy model consisting of two tanh transitions allowing for an additional high-redshift ionization plateau. We demonstrated accuracy of RELIKE for both examples by comparing model parameter constraints with those obtained from sampling the exact Planck likelihoods using a modified CAMB code. We also showed excellent agreement between the confidence levels of the cumulative optical depth evolution between those of the exact ionization models and their 5-PC representations.

Besides conducting explicit model testing with RELIKE, we also extracted model-independent constraints using the PC chains themselves. We summarize these results below.
(i) We obtained the total optical depth constraint $\tau_{PC} = 0.0619^{+0.0086}_{-0.0068}$ (68% C.L.) using PCs in our default likelihoods $\text{plik\_lite\_TTTEEE + lowl + SRoll2}$.

(ii) Using the $\text{SRoll2}$ likelihood tightened the PC constraint on $\tau$ and shifted it toward slightly higher values of $\tau$ compared to the $\text{plik\_lite\_TTTEEE + lowl + SimAll}$ result $\tau_{PC} = 0.0582^{+0.0072}_{-0.0080}$ (68% C.L.).

(iii) We checked the robustness of our $z_{\text{max}} = 30$ results by performing a PC analysis of the Planck data with $z_{\text{max}} = 50$ where 7 PCs were used for the larger redshift range. The results changed negligibly: $\tau_{PC} = 0.0626^{+0.0061}_{-0.0072}$ (68% C.L.).

(iv) The constraint on the high-redshift optical depth $\tau_{PC}(15,30) < 0.020$ (95% C.L.) is also stable when extending $z_{\text{max}}$ to 50: $\tau_{PC}(15,50) < 0.019$ (95% C.L.).

(v) Our high-redshift optical depth upper limit is however a factor of $\sim 3$ larger than that from the Planck 2018 cosmological parameter paper $\tau(15,30) < 0.007$ (95% C.L.) obtained from another model-independent method called FlexKnot. Using explicitly a two-step toy model (without using any PCs), we tested that $\tau_{2\text{step}}(15,30) < 0.016$ (95% C.L.) is still allowed, which is consistent with our PC results.

Finally, we clarified the effects of applying a range-bound prior on the PC amplitude while deriving model-independent constraints on the optical depth (note that this prior does not enter the $\text{RELIKE}$ code). We demonstrated (see Fig. 12) that removing more unphysical models would only loosen the high-$z$ optical depth upper limit, in contrary to what was claimed in Ref. [17].

Given that the CMB intrinsically constrains integrated quantities of the ionization history, we recommend that model-independent constraints be formulated in the cumulative optical depth space rather than the ionization history space. There is also a tendency in the literature to compare a specific model with more general constraints on $x_\text{s}(z)$. We caution the reader that these constraints are derived with a specific prior assumption and sometimes on a limited set of models. So our recommendation for constraining specific models is to use the $\text{RELIKE}$ code which properly returns the effective Planck likelihood for any specific $x_\text{s}(z)$ function.

We also recommend using $\text{RELIKE}$ to properly combine CMB large-scale datasets with other reionization datasets such as kSZ, galaxy luminosity functions, quasars spectra, line-intensity mapping, star formation rate, etc. In the past, joint analyses of the reionization history often used the Planck tanh constraint as a $\tau$ prior, which is an approximate proxy for the actual likelihood of the model in question. With the release of $\text{RELIKE}$, one can now easily obtain the full Planck constraint on the specific model and conduct consistent joint analyses.

To do so, one can interface the python $\text{RELIKE}$ package with existing MCMC samplers for generic data such as Cobaya and CosmoSIS, or those adapted for specific reionization datasets (e.g., 21 CMMC [32] and CosmoMCREion [33]). When the joint dataset is constraining enough (e.g., kSZ), one can sample while varying cosmological parameters since the $\text{RELIKE}$ likelihood is valid for cosmologies consistent with the Planck best-fit. We expect that joint analyses work in the future (see e.g., Ref. [30,34] for some recent examples) would be made much easier with $\text{RELIKE}$.

In the future, $\text{RELIKE}$ can be extended in several ways. Recent studies of quasars may indicate that the Universe have not been entirely ionized by $z = 6$ ($\sim 80\%$ ionized at $z = 6$) [35]. If this result is confirmed, the PC approach can be easily modified for a smaller $z_{\text{min}}$ in a manner similar to the $z_{\text{max}}$ study here.

Likewise current and future CMB experiments such as CLASS [36,37] and LiteBIRD [38] may also improve on Planck and eventually provide cosmic variance limited measurements of CMB polarization on the largest scales, yielding improved reionization constraints. The same approach used here can be easily updated to capture the better constraints from these data in a joint effective likelihood.

In sum, our release of the $\text{RELIKE}$ code significantly reduces the time to constrain specific ionization models with Planck. It also opens the way for consistent and fast joint analyses of current and future CMB large-scale polarization data with a variety of increasingly rich reionization datasets in the next decades.

ACKNOWLEDGMENTS

We thank Tzu-Ching Chang, GuoChao Sun, Lluis Mas-Aribas and Olivier Doré for useful discussions. W. H. was supported by U.S. Dept. of Energy Contract No. DE-FG02-13ER41958 and the Simons Foundation.
[1] N. Aghanim et al. (Planck Collaboration), Astron. Astrophys. 641, A6 (2020).
[2] Understanding the Epoch of Cosmic Reionization: Challenges and Progress, edited by A. Mesinger, Astrophysics and Space Science Library, Vol. 423 (Springer, New York, 2016).
[3] K. M. Smith, W. Hu, and M. Kaplinghat, Phys. Rev. D 74, 123002 (2006).
[4] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, Phys. Rev. D 92, 123535 (2015).
[5] W. Hu and B. Jain, Phys. Rev. D 70, 043009 (2004).
[6] R. A. Sunyaev and Y. B. Zeldovich, Astrophys. Space Sci. 7, 3 (1970).
[7] R. A. Sunyaev and Y. B. Zeldovich, Annu. Rev. Astron. Astrophys. 18, 537 (1980).
[8] M. McQuinn, S. R. Furlanetto, L. Hernquist, O. Zahn, and M. Zaldarriaga, Astrophys. J. 630, 643 (2005).
[9] A. Mesinger, M. McQuinn, and D. N. Spergel, Mon. Not. R. Astron. Soc. 422, 1403 (2012).
[10] W. Hu and G. P. Holder, Phys. Rev. D 68, 023001 (2003).
[11] M. J. Mortonson and W. Hu, Astrophys. J. 672, 737 (2008).
[12] M. J. Mortonson and W. Hu, Astrophys. J. 686, L53 (2008).
[13] C. H. Heinrich, V. Miranda, and W. Hu, Phys. Rev. D 95, 023513 (2017).
[14] P. A. R. Ade et al. (Planck Collaboration), Astron. Astrophys. 571, A22 (2014).
[15] G. Obied, C. Dvorkin, C. Heinrich, W. Hu, and V. Miranda, Phys. Rev. D 98, 043518 (2018).
[16] W.-M. Dai, Z.-K. Guo, and R.-G. Cai, Phys. Rev. D 92, 123521 (2015).
[17] M. Millea and F. Bouchet, Astron. Astrophys. 617, A96 (2018).
[18] C. Heinrich and W. Hu, Phys. Rev. D 98, 063514 (2018).
[19] J. M. Delouis, L. Pagano, S. Mottet, J. L. Puget, and L. Vibert, Astron. Astrophys. 629, A38 (2019).
[20] G. D. Becker, J. S. Bolton, and A. Lidz, Pub. Astron. Soc. Aust. 32, e045 (2015).
[21] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. 538, 473 (2000).
[22] C. Howlett, A. Lewis, A. Hall, and A. Challinor, J. Cosmol. Astropart. Phys. 04 (2012) 027.
[23] G. D. Becker, J. S. Bolton, M. G. Haehnelt, and W. L. W. Sargent, Mon. Not. R. Astron. Soc. 410, 1096 (2011).
[24] A. Lewis, Phys. Rev. D 87, 103529 (2013).
[25] A. Lewis and S. Bridle, Phys. Rev. D 66, 103511 (2002).
[26] N. Aghanim et al. (Planck Collaboration), Astron. Astrophys. 641, A5 (2020).
[27] J. Torrado and A. Lewis, J. Cosmol. Astropart. Phys. 05 (2021) 057.
[28] J. Zuntz, M. Paterno, E. Jennings, D. Rudd, A. Manzotti, S. Dodelson, S. Bridle, S. Sehrish, and J. Kowalkowski, Astron. Comput. 12, 45 (2015).
[29] L. Pagano, J. M. Delouis, S. Mottet, J. L. Puget, and L. Vibert, Astron. Astrophys. 635, A99 (2020).
[30] K. Ahn and P. R. Shapiro, Astrophys. J. 914, 44 (2021).
[31] D. Paolletti, D. K. Hazra, F. Finelli, and G. F. Smoot, J. Cosmol. Astropart. Phys. 09 (2020) 005.
[32] B. Greig and A. Mesinger, Mon. Not. R. Astron. Soc. 449, 4246 (2015).
[33] A. Chatterjee, T. R. Choudhury, and S. Mitra, arXiv:2101.11088.
[34] Y. Qin, V. Poulin, A. Mesinger, B. Greig, S. Murray, and J. Park, Mon. Not. R. Astron. Soc. 499, 550 (2020).
[35] G. D. Becker, A. D’aloisio, H. M. Christenson, Y. Zhu, G. Worseck, and J. S. Bolton, arXiv:2103.16610.
[36] D. J. Watts et al., Astrophys. J. 863, 121 (2018).
[37] D. J. Watts, G. E. Addison, C. L. Bennett, and J. L. Weiland, 10.3847/1538-4357/ab5fd5, arXiv:1910.00590.
[38] M. Hazumi et al. (LiteBIRD Collaboration), Proc. SPIE Int. Soc. Opt. Eng. 11443, 114432F (2020).