Complete reionization constraints from Planck 2015 polarization

Chen He Heinrich, Vinicius Miranda, and Wayne Hu

Complete Reionization Constraints from Planck 2015 Polarization

Chen He Heinrich,[1, 2] Vinicius Miranda,[3] and Wayne Hu[1, 4]

[1] *Kavli Institute for Cosmological Physics, Enrico Fermi Institute, University of Chicago, Chicago IL 60637*
[2] *Department of Physics, University of Chicago, Chicago IL 60637*
[3] *Center for Particle Cosmology, Department of Physics and Astronomy,*
University of Pennsylvania, Philadelphia, PA 19104
[4] *Department of Astronomy & Astrophysics, University of Chicago, Chicago IL 60637*

We conduct an analysis of the Planck 2015 data that is complete in reionization observables from the large angle polarization E-mode spectrum in the redshift range $6 < z < 30$. Based on 5 principal components, all of which are constrained by the data, this single analysis can be used to infer constraints on any model for reionization in the same range; we develop an effective likelihood approach for applying these constraints to models. By allowing for an arbitrary ionization history, this technique tests the robustness of inferences on the total optical depth from the usual step-like transition assumption, which is important for the interpretation of many other cosmological parameters such as the dark energy and neutrino mass. The Planck 2015 data not only allow a high redshift $z > 15$ component to the optical depth but prefer it at the 2σ level. This preference is associated with excess power in the multipole range $10 \lesssim \ell \lesssim 20$ and may indicate high redshift ionization sources or unaccounted for systematics and foregrounds in the 2015 data.

I. INTRODUCTION

The epoch of reionization of the universe remains one of the least well understood aspects of the standard model of cosmology (see e.g. [1]). Yet its impact on the interpretation of its fundamental properties is comparatively large in this era of precision cosmology. In addition to the intrinsic astrophysical interest in ionization sources, reionization uncertainties impact CMB inferences on the initial power spectrum and hence also cosmic acceleration through the growth of structure [2]. In the future it will be one of the leading sources of error in the interpretation of neutrino mass measurements from gravitational lensing [3, 4], the study of large scale anomalies in the CMB [5, 6], and the inflationary consistency relation [7].

The standard approach to parameterizing the impact of reionization on the CMB is with the total Thomson optical depth through the reionization epoch. Although it is indeed the total optical depth that is important for the interpretation of most other aspects of cosmology, it is usually assumed that reionization occurs promptly in a step-like transition. Interestingly, the central value of the inferred optical depth from this approach has been steadily drifting downwards from its first detection in WMAP1 [8] to the current but still proprietary Planck 2016 HFI results [9, 10].

Relaxing this sharp transition assumption can in principle raise the optical depth inference from the CMB as well as change its implications for sources of high redshift ionization (e.g. [11]). In particular, the angular scale of the peak and the width of the reionization bump in the E-mode CMB polarization power spectrum carries coarse grained information on the redshift dependence of the ionization history.

There is an alternate, model independent approach introduced in Ref. [12] that fully addresses these concerns. The impact of *any* ionization history on the large angle CMB polarization spectrum can be completely characterized by a handful of reionization parameters constructed from the principal components (PCs) of the ionization history with respect to the E-mode power spectrum. This approach has the advantage over redshift binning alternatives of being observationally complete without introducing numerous highly correlated parameters [13]. Conversely, it does not provide an accurate, local reconstruction for visualizing the ionization history itself.

This approach was implemented and tested on WMAP3 [14] and WMAP5 [15] power spectra which showed that those data allowed for, but did not particularly favor, contributions to the optical depth from high redshift. It was adopted in the Planck 2013 analysis but exclusively to test the impact of marginalizing the ionization history on inflationary parameters rather than drawing inferences on reionization itself [16]. The impact on massive neutrinos and gravitational wave inferences was also examined in Ref. [17].

In this work, we analyze the public Planck 2015 data, including the LFI large angle polarization power spectrum, using the observationally complete PC basis. In addition we further develop the technique as a method to probe reionization itself. This development is timely as the technique should come to its full fruition with the upcoming final release of Planck data which will be the definitive result on large angle polarization for years to come.

We demonstrate that the Planck 2015 data already have more information on the ionization history than just the total optical depth and provide effective likelihood tools for interpreting this information in any given model for reionization within the redshift range analyzed. Tested here, these techniques can be straightforwardly applied to the final release when it becomes available.

We begin with a review of the approach itself in §II. In §III, we analyze the Planck 2015 data and explore the origin of the additional information on the high redshift

ionization history. We develop and test an effective likelihood approach for utilizing our analysis to constrain the parameters of any given model of reionization from $6 < z < 30$ in §IV. We discuss these results in §V.

II. COMPLETE REIONIZATION BASIS

We briefly summarize the principal component technique for the complete characterization of reionization constraints from the large angle C_ℓ^{EE} polarization spectrum as introduced in Ref. [12] and implemented in [14, 15].

We parametrize the ionization fraction relative to fully ionized hydrogen $x_e(z)$ into its principal components with respect to the E-mode polarization of the CMB [12]:

$$x_e(z) = x_e^{\text{fid}}(z) + \sum_a m_a S_a(z). \qquad (1)$$

Here m_a are the PC amplitudes and $S_a(z)$ are the eigenfunctions of the Fisher information matrix for $x_e(z)$ in a given range $z_{\min} < z < z_{\max}$ from cosmic variance limited C_ℓ^{EE} measurements, and $x_e^{\text{fid}}(z)$ is the fiducial model around which the Fisher matrix is computed. In practice, we discretize the redshift space to $\delta z = 0.25$, well beyond the resolution limit of CMB observables, and assume linear interpolation between points to form the continuous functions $S_a(z)$. The components are rank ordered by their Fisher-estimated variances and in practice the first 5 components carry all the information in C_ℓ^{EE} to the cosmic variance limit [12]. In this work, we therefore truncate the PC expansion and retain 5 m_a parameters to describe reionization. We take $z_{\min} = 6$ to be consistent with Lyα forest constraints (e.g. [18]) and $z_{\max} = 30$.

In this truncated representation, the 5 PC decomposition is *not* complete in the ionization history itself. Instead it is a complete representation of the *observable* impact on C_ℓ^{EE} of any given $x_e^{\text{true}}(z)$ through its projection onto the 5 PC basis

$$m_a = \int_{z_{\min}}^{z_{\max}} dz \, \frac{S_a(z)[x_e^{\text{true}}(z) - x_e^{\text{fid}}(z)]}{z_{\max} - z_{\min}}. \qquad (2)$$

When reconstructed through Eq. (1) with 5 PCs, $x_e(z) \neq x_e^{\text{true}}(z)$, even though it models the observables to high precision. The PC analysis therefore is a forward tool to infer constraints on all possible ionization histories between $z_{\min} < z < z_{\max}$ with a single analysis, not an inverse tool that reconstructs the ionization history.

We use a modified version of CAMB[1] [19] to compute the CMB power spectra given the ionization history. Because CAMB integrates the Boltzmann equation
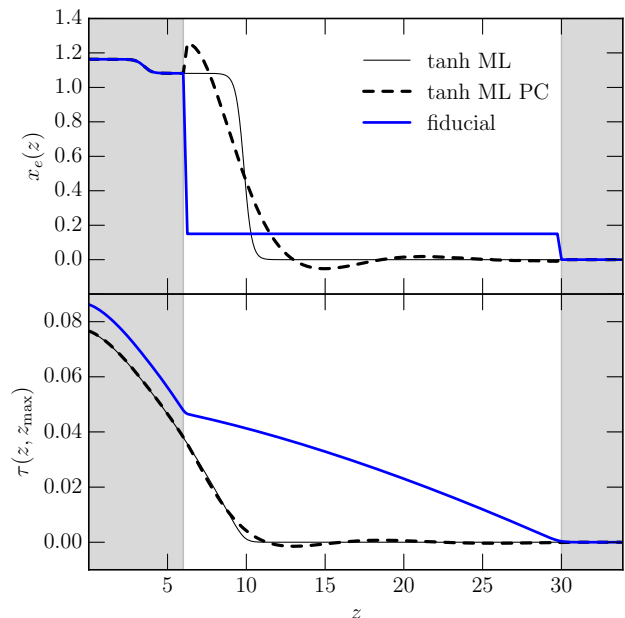
FIG. 1. Tanh reionization (black thin) vs. fiducial model (blue thick) around which the PC decomposition (in the unshaded region) is centered. *Top*: ionization fraction $x_e(z)$ as a function of redshift. For the tanh model, despite modeling the polarization observables to high accuracy, the x_e reconstruction from 5 PCs is poor (black dashed). *Bottom*: cumulative optical depth from z to $z_{\max}$. The tanh model corresponds to the maximum likelihood in the MCMC chain of §III and with its sudden reionization accumulates its optical depth exclusively at low redshift in contrast with the fiducial model. The 5 PC reconstruction is a better approximation for $\tau(z, z_{\max})$, especially for the difference between low and high z.

by parts, it requires a smooth ionization history for numerical stability, whereas the $S_a(z)$ are continuous but not smooth. Consequently we smooth the ionization history in Eq. (1) with a Gaussian in $\ln(1 + z)$ of width $\sigma_{\ln(1+z)} = 0.015$. This does not affect our results in a statistically significant way. However for consistency, when integrating the ionization history to form the cumulative Thomson optical depth

$$\tau(z, z_{\max}) = n_{\text{H}}(0)\sigma_T \int_z^{z_{\max}} dz \, \frac{x_e(z)(1 + z)^2}{H(z)}, \qquad (3)$$

we employ the smoothed x_e which formally has support beyond the bounds. We include this small correction by integrating slightly past $z_{\max}$ in practice. Here $n_{\text{H}}(0)$ is the hydrogen number density at $z = 0$.

For the fiducial ionization history, we take $x_e^{\text{fid}} = 0.15$ for $6 < z < 30$ on the $\delta z = 0.25$ spaced discrete points in the interval with linear interpolation in between, in order to let the PC amplitudes fluctuate the ionization history both up and down without entering the unphysical region $x_e < 0$, where the number is chosen to give a reasonable $\tau(0, z_{\max})$. It is important to note that the PCs allow arbitrarily large deviations from the fiducial model, and so this choice does not bias results. For $z \geq 30$ we assume x_e follows the ionization history from recombination. For
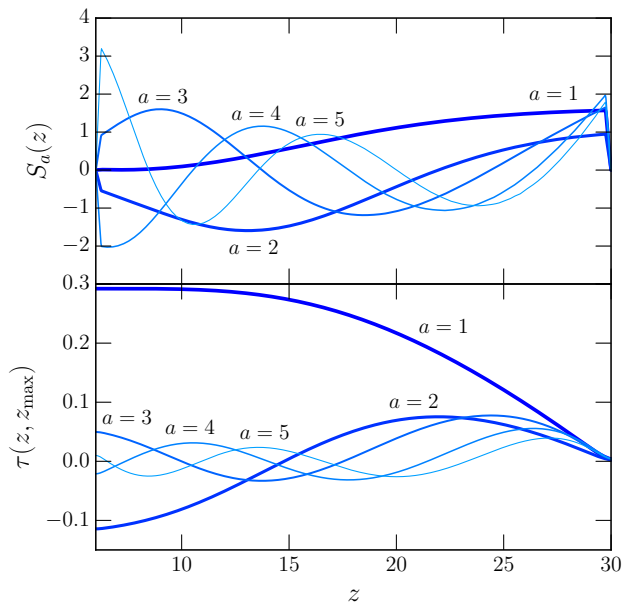
FIG. 2. First five PCs forming a complete set of polarization observables for the redshift range displayed. *Top*: ionization PCs in redshift. PCs are continuous functions with linear interpolation between discrete values at $\delta z = 0.25$ with boundaries at z_{min} and z_{max} held fixed to zero. *Bottom*: cumulative optical depth for unit amplitude modes as in Fig. 1. The first mode ($a = 1$) mainly changes the total optical depth $\tau(0, z_{max})$ from the fiducial model. The second mode ($a = 2$) adjusts the low and high redshift optical depth contributions whereas the higher modes make finer and less observable modifications to its redshift distribution.

$z \leq 6$ we assume fully ionized hydrogen and singly ionized helium

$$x_e = 1 + f_{He}, \qquad (4)$$

for $z \gtrsim z_{He}$ and doubly ionized helium

$$x_e = 1 + 2f_{He}, \qquad (5)$$

for $z \lesssim z_{He}$, where

$$f_{He} = \frac{n_{He}}{n_H} = \frac{m_H}{m_{He}} \frac{Y_p}{1 - Y_p} \qquad (6)$$

is the ratio of the helium to hydrogen number density. We take the helium mass fraction Y_p to be consistent with big bang nucleosynthesis given the baryon density. Following CAMB, we take this helium reionization transition to be mediated by a tanh function in redshift centered at $z_{He} = 3.5$ [20] with width $\Delta z = 0.5$.

In Fig. 1, we show the fiducial ionization history (thick blue) and contrast it with the standard approach of CAMB (thin black) that takes hydrogen and singly ionized helium reionization to be given by the tanh form

$$x_e^{true}(z) = \frac{1 + f_{He}}{2} \left\{ 1 + \tanh\left[\frac{y(z_*) - y(z)}{\Delta y} \right] \right\}, \quad (7)$$

with $y(z) = (1 + z)^{3/2}$, $\Delta y = (3/2)(1 + z)^{1/2}\Delta z$, and $\Delta z = 0.5$. We take here $z_* = 9.85$, corresponding the

chain maximum likelihood model ($\tau = 0.0765$) from §III, for illustrative purposes. Projected onto 5 PCs and re-summed into $x_e(z)$, Eq. (1) yields a poor reconstruction of the ionization history itself. Nonetheless as we shall see in Fig. 9, the PC decomposition provides an excellent representation of the polarization power spectrum.

We also show in Fig. 1 the cumulative optical depth from z to z_{max}. Although the two models have comparable total optical depth $\tau(0, z_{max})$ the fiducial model receives much of its contribution from high redshift. Note also that although we do not allow for uncertainties in helium reionization, its entire impact is a small correction on an already small contribution to τ. Furthermore, although the reconstruction of $\tau(z, z_{max})$ using 5 PCs is still imperfect (black dashed line), it is much better than $x_e(z)$ as it is more closely related to the polarization observables. The PC reconstruction smooths out sharp transitions in the cumulative optical depth but gives an accurate representation of the high and low redshift contributions of the model. In the analysis below, in order to compare exactly the same statistic $\tau(z, z_{max})$ between models, we always employ the PC reconstructed version.

In Fig. 2 we show the 5 PC ionization functions $S_a(z)$ which allow observationally complete variations around the fiducial model. We also show the cumulative optical depth $\tau(z, z_{max})$ for a unit amplitude m_a in each mode. The lowest-variance eigenmode S_1 adjusts the total optical depth, mainly from the high redshift end. The S_2 mode allows a redistribution of the optical depth between high redshift and low redshift. The higher modes allow finer adjustments in the redshift distribution of the optical depth and carry very little total optical depth $\tau(0, z_{max})$.

For the Planck data set, most of the information in the ionization history is carried by the first two modes and therefore relates to the amount of high vs. low redshift optical depth. We keep all 5 PCs for completeness in representing the observable impact of a given ionization history and to marginalize uncertainties that they introduce.

In fact, PCs with no prior constraints on the mode amplitudes m_a allow deviations that are unphysical $x_e < 0$ and $x_e > x_e^{max}$. With a truncation at 5 PCs, physicality cannot be strictly enforced since the missing eigenmodes, while irrelevant for the observables, can restore physicality of a model. We follow Ref. [15] in placing necessary but not sufficient conditions for physicality

$$\sum_{a=1}^{5} m_a^2 \leq (x_e^{max} - x_e^{fid})^2, \qquad (8)$$

where $x_e^{fid} = 0.15$ and $m_a^- \leq m_a \leq m_a^+$ with

$$m_a^{\pm} = \int_{z_{min}}^{z_{max}} dz \frac{S_a(z)[x_e^{max} - 2x_e^{fid}(z)] \pm x_e^{max}|S_a(z)|}{2(z_{max} - z_{min})}. \qquad (9)$$

In principle $x_e^{max} = 1 + 2f_{He}$ but in practice for compatibility with the ionization state at z_{min} we take $x_e^{max} =$
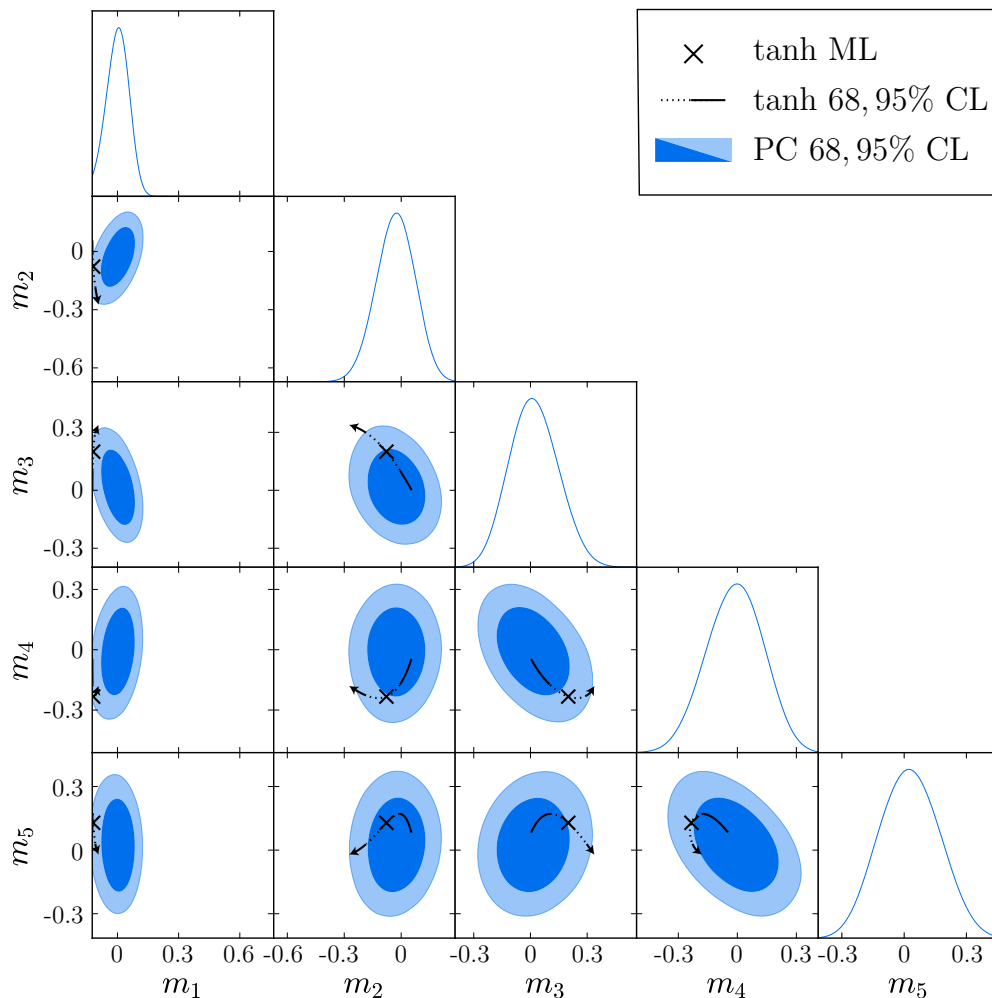
FIG. 3. Reionization constraints from Planck 2015 data on PC amplitudes compared with those of the tanh model projected onto PCs. 1D posterior probability distributions are shown on the upper diagonal and 2D 68% and 95% CL regions in the $m_a - m_b$ planes in the lower triangle. Tanh constraints are shown as trajectories in these planes (68% CL dotted lines; 95% CL solid lines) with the arrow pointed to higher τ. The maximum likelihood ($\times$) and allowed region for the tanh model is disfavored in the observationally complete PC parameter space. Box boundaries represent physicality priors on the ionization history.

$1 + f_{\mathrm{He}}$. Since this prior is applied very conservatively as a necessary condition, we in fact still allow ionization fractions out beyond the true maximum with this prior.

For analyzing any physical model of hydrogen reionization, $x_e^{\mathrm{true}}(z)$, using our 5 PC decomposition from Eq. (2), the priors are noninformative as they are automatically satisfied. We apply them mainly for visualizing whether the data constrain the mode amplitudes significantly better than the physicality bounds.

III. REIONIZATION CONSTRAINTS

We use the Markov Chain Monte Carlo (MCMC) technique to sample from the posterior probability density in the reionization and cosmological parameter space. Our main analysis is on the Planck 2015 data [21] using the public likelihoods plik_lite_TTTEEE for high-ℓ's and

lowTEB for low-ℓ's, which includes LFI but not HFI polarization.[2] For our cosmological parameters, we vary the base set of ΛCDM parameters (baryon density $\Omega_b h^2$, cold dark matter density $\Omega_c h^2$, effective acoustic scale θ_{MCMC}, scalar power spectrum amplitude $\ln(10^{10} A_s)$ and tilt n_s). We fix the neutrinos to their minimal contribution of one massive species with $m_\nu = 0.06 \mathrm{eV}$. To those parameters we add the 5 PC mode amplitudes $m_1, \ldots, m_5$. For comparison we also run a separate chain with the standard tanh ionization history, parametrized by the total reionization optical depth τ. We assume flat priors in each of

---

[2] We have tested that our results are robust to explicitly marginalizing foreground parameters, as opposed to using the premarginalized likelihood, by separately running a PC MCMC using the lowTEB low-ℓ likelihood and the plikHM high-ℓ likelihood.
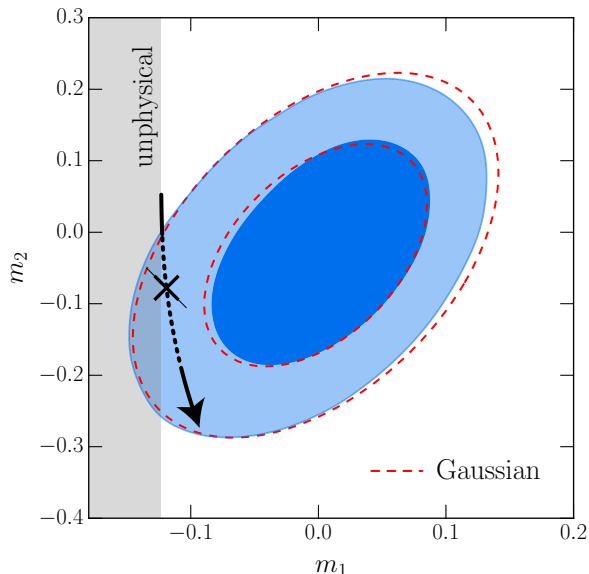
FIG. 4. Best constrained PC plane, $m_1 - m_2$ (magnified from Fig. 3). Tanh trajectories pass near the boundary of the unphysical region (gray shaded) at low τ since reionization must be complete by $z = 6$. Also shown is the multivariate Gaussian approximation to the PC posterior using the means and covariance from Tab. I (red dashed lines, 68% and 95% CL, see Eq. 16).
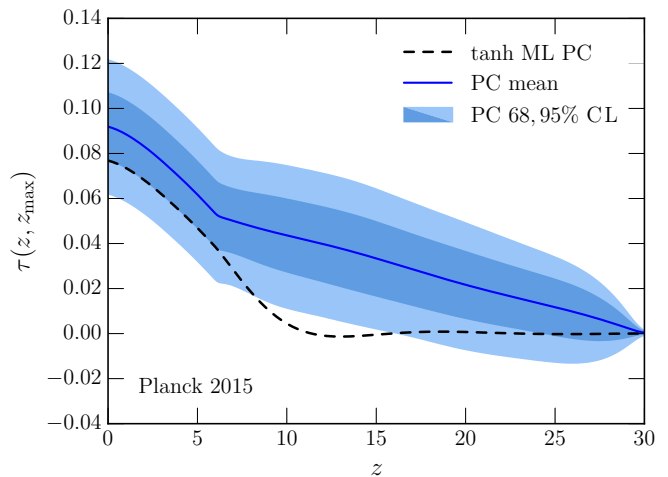

FIG. 5. Cumulative optical depth $\tau(z, z_{\max})$ in the Planck 2015 analysis. 68% and 95% CL constraints from the complete PC analysis compared with the chain ML tanh model also constructed from PCs (dashed line). PC analysis show that the data prefer finite high redshift optical depth out to $z \gtrsim 16$ at 95% CL. The tanh model has essentially no optical depth for $z \gtrsim 10$ due to its step-like form. Both model classes have the same ionization history for $z < 6$.

the given parameters.

In Fig. 3, we show the 1D and 2D marginalized posteriors in the PC amplitudes m_a and for comparison the standard tanh ionization history with 68% and 95% ranges in its parameter τ projected onto the PCs. Box bounds represent the physicality prior from Eq. (9). In Tab. I we give the corresponding means $\bar{m}_a$, errors $\sigma(m_a)$ and correlation matrix R_{ab} which define the covariance matrix C_{ab} as

$$C_{ab} = \sigma(m_a)\sigma(m_b)R_{ab}. \tag{10}$$

Note that although the PCs are constructed to be uncorrelated for infinitesimal deviations from the fiducial model, they do not remain so for finite deviations (see [15]).

Although all 5 PCs are measured to better than their physicality priors, unlike in the WMAP5 analysis of Ref. [15], only m_1 and m_2 are bounded substantially better. In Fig. 4 we highlight this plane. Note that in the tanh model, low m_1 and high m_2 corresponds to small

TABLE I. PC chain means $\bar{m}_a$, standard deviations $\sigma(m_a)$, and correlation matrix R_{ab}.

	$\bar{m}_a$	$\sigma(m_a)$	m_1	m_2	m_3	m_4	m_5
m_1	0.002	0.053	1.000	0.450	−0.432	0.273	−0.073
m_2	−0.030	0.101	0.450	1.000	−0.262	0.055	0.072
m_3	0.019	0.128	−0.432	−0.262	1.000	−0.417	0.155
m_4	−0.012	0.143	0.273	0.055	−0.417	1.000	−0.428
m_5	0.026	0.143	−0.073	0.072	0.155	−0.428	1.000

values of the total optical depth and so these models skirt the edge of physicality there given that the universe must be reionized by $z = 6$. In this MCMC, physicality priors are actually imposed after the fact by eliminating samples in the shaded region. Likewise the covariance matrix in Eq. (10), is calculated before these samples are removed. This ensures that the Gaussian approximation to the posterior is not distorted by the prior (see Eq. 16). As shown in Fig. 4, the Gaussian approximation is fairly good in the 68% and 95% CL regions of the $m_1 - m_2$ plane, and we have checked that it is equally good in the other planes.

Interestingly, the standard tanh model is moderately disfavored in the wider, observationally complete, PC space leading to changes in the inference for the optical depth summarized in Tab. II. In Fig. 6, we compare the posterior probability distributions for the total $\tau(0, z_{\max})$ between the two chains. In the PC analysis the total optical depth constraints shift up by almost 1σ from the tanh analysis while having comparable widths. In ΛCDM the main consequence is that the amplitude of structure at $z = 0$ goes from $\sigma_8 = 0.831 \pm 0.013$ in the tanh model to $\sigma_8 = 0.840 \pm 0.012$ in the PC analysis.

These changes are related to the fact that the standard tanh models only skirt just inside the 2D 95% CL even near the maximum likelihood found in the chain (see Fig. 4). The tanh ML model have the following parameters: $\Omega_b h^2 = 0.02224$, $\Omega_c h^2 = 0.1197$, $\theta_{\mathrm{MCMC}} = 1.04075$, $\ln(10^{10} A_s) = 3.0866$, $n_s = 0.9653$ and $\tau = 0.0765$. Projected onto the PC space through Eq. (2), this value of τ corresponds to the m_a parameter vector

$$\mathbf{m}^T = \{-0.119, -0.078, 0.200, -0.233, 0.129\}. \tag{11}$$

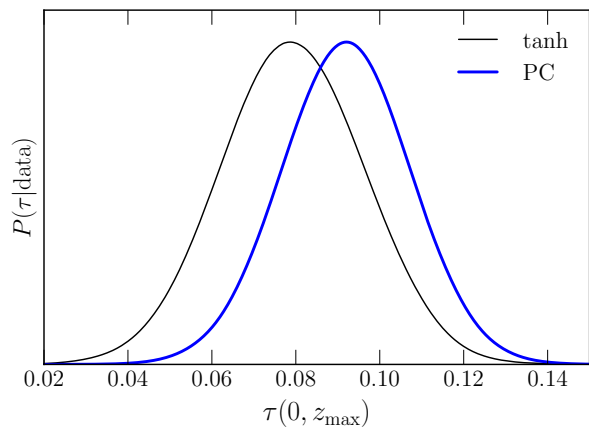Note that for the single best constrained component m_1,

FIG. 6. Posterior probability distribution of the total optical depth from the Planck 2015 tanh vs. PC analyses. The mean of the PC distribution is approximately 1σ higher than the tanh distribution while the width remains comparable.

the tanh ML model is more than 2σ off the mean.

In Fig. 5, we explore the physical origin of this difference. Here we show the 68% and 95% CL constraints on the cumulative optical depth $\tau(z, z_{\max})$ as a parameter derived from the m_a posterior probability. The shape for $z < 6$ is fixed by assumption. Note that at $z \gg 1$, the cumulative optical depth from Eq. (3) becomes approximately

$$\tau(z, z_{\max}) \propto \frac{\Omega_b h^2 (1 - Y_p)}{(\Omega_m h^2)^{1/2}} \int_z^{z_{\max}} dz\, x_e(z)(1+z)^{1/2},$$
(12)

where $\Omega_m h^2$ parametrizes the sum over all of the non-relativistic density components. Integrals over redshift in Eq. (3) can be precomputed for individual PCs once and for all in a fiducial cosmology as in Fig. 2 and then just summed with m_a weights and a rescaled prefactor following Eq. (12).

At the 95% CL, the data favor optical depth contributions at $z \gtrsim 16$. For comparison, we also plot the ML tanh model, projected onto the PC basis and calculated in the same way. This model has essentially no optical depth contributions for $z > 10$, not because the data forbid it but because of the functional form of the model. More generally, because of its step-like form, the tanh family of models cannot generate high redshift optical depth without also overproducing the total optical depth $\tau(0, z_{\max})$ (see Tab. II).

Given these differences, we explore further their origin in the data. Constraints on the optical depth are also affected by the temperature power spectrum indirectly and directly. Gravitational lens effects place constraints on the amplitude of the matter power spectrum through A_s and so in combination with measurements of the temperature power spectrum which determine $A_s e^{-2\tau(0,z_{\max})}$ constrain $\tau(0, z_{\max})$ indirectly [22]. In particular, the Planck temperature power spectrum favors more gravitational lensing than is predicted by the best fit ΛCDM
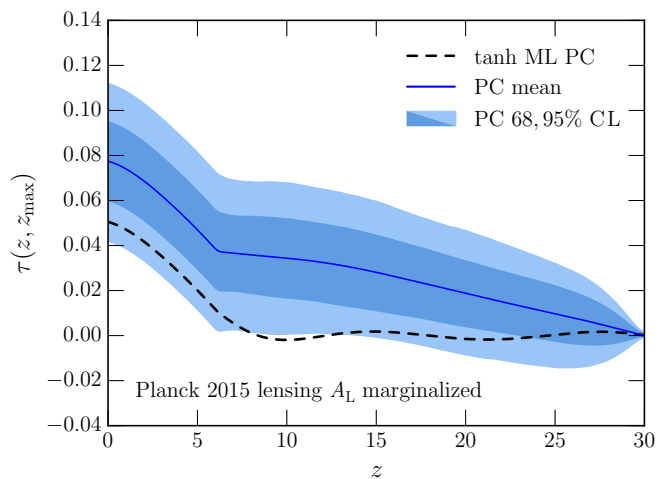

FIG. 7. Cumulative optical depth $\tau(z, z_{\max})$ after marginalizing the lensing amplitude A_L in the Planck 2015 analysis (as in Fig. 5 otherwise). The total optical depth decreases for both the PC and tanh analyses but mainly by lowering the low redshift contributions. Preference for a finite high redshift contribution $\tau(15, z_{\max})$ remains at near the 95% CL.

parameters and hence tends to drive τ to larger values [23].

To test whether the preference for high redshift optical depth originates from gravitational lensing and not large angle polarization, we follow the Planck 2015 analysis and marginalize a multiplicative renormalization of the lens power spectrum A_L by adding it to the parameter set of a new MCMC analysis of both the PC and the tanh model. In the PC case we obtain $A_L = 1.11 \pm 0.07$ and in the tanh case $A_L = 1.15 \pm 0.08$. In Fig. 7, we show the impact on $\tau(z, z_{\max})$. As expected the total optical depth $\tau(0, z_{\max})$ is approximately 1σ lower but notably the high redshift preference weakens only moderately. A finite $\tau(15, z_{\max})$ is still favored at nearly the 95% CL (see also Tab. II). We conclude that much of the preference for high vs. low redshift optical depth comes from the large angle polarization data itself.

To further test this conclusion, we replace the Planck polarization data with WMAP9. In order to consistently analyze WMAP9 and Planck data sets with publicly available likelihood codes, we also employ Planck 2013 instead of 2015 data. Fig. 8 shows that this replacement has a larger impact on the high redshift end with the preference for $\tau(15, z_{\max})$ dropping to almost 1σ (see also Tab. II). Conversely, the constraints on m_3, m_4, m_5 remain largely the same as the baseline Planck 2015 analysis indicating that they are constrained mainly by the temperature data. As noted in Ref. [15] the Doppler effect from reionization imprints features on the temperature spectrum which constrain the higher order PCs. With the extended reach of the Planck temperature power spectrum as compared with WMAP5, these features can be better separated from cosmological parameters.
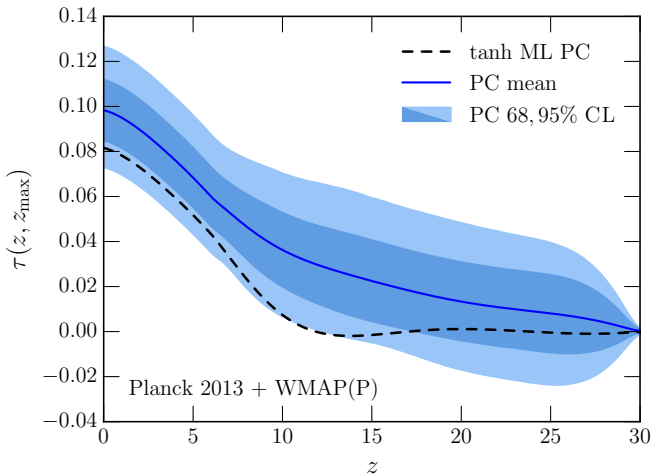
FIG. 8. Cumulative optical depth $\tau(z, z_{\max})$ in the Planck 2013 analysis with WMAP9 polarization (as in Fig. 5 otherwise). The total optical depth increases slightly for both the PC and tanh analyses while the high redshift contributions decrease. Preference for a finite high redshift contribution $\tau(15, z_{\max})$ is lowered to closer to 68% CL.
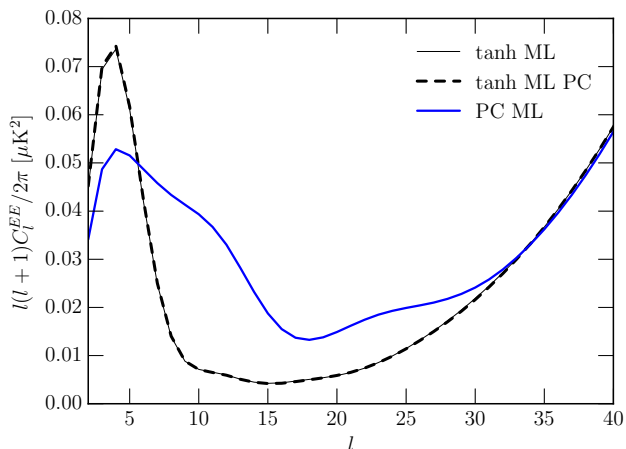

FIG. 9. Polarization power spectrum C_l^{EE} in the maximum likelihood models from the Planck 2015 PC and tanh chains. Preference for high redshift optical depth in the PC analysis corresponds to a power spectrum with a broader reionization bump extending to higher multipoles than the tanh model. To fit $\ell \lesssim 10$ data, tanh models predict much less power at $10 \lesssim \ell \lesssim 20$. Also shown is the tanh ML model calculated from the projection onto the 5 PC basis, showing its high accuracy and observational completeness. Note that the chain maximum is not necessarily the true maximum.

Since the preference for high redshift ionization mainly originates from the low multipole polarization data, it is interesting to compare C_ℓ^{EE} for the maximum likelihood tanh model from Eq. (11) to the maximum likelihood PC parameters. The PC chain maximum has $\Omega_b h^2 = 0.02230$, $\Omega_c h^2 = 0.1195$, $\theta_{\mathrm{MCMC}} = 1.04078$, $\ln(10^{10} A_s) = 3.1163$, $n_s = 0.9659$ and

$$\mathbf{m}^T = \{0.029, 0.009, -0.026, 0.085, 0.062\}. \tag{13}$$

Note that in both cases the ML is simply the maxi-

TABLE II. Total and high redshift optical depth constraints for different model and data sets combinations. Tanh models allow negligible high redshift optical depth whereas finite values are favored in the PC space.

Model	Data	$\tau(0, z_{\max})$	$\tau(15, z_{\max})$
PC	P15	0.092 ± 0.015	0.033 ± 0.016
tanh	P15	0.079 ± 0.017	–
PC $+A_L$	P15	0.078 ± 0.018	0.028 ± 0.016
tanh $+A_L$	P15	0.056 ± 0.020	–
PC	P13+WMAP(P)	0.098 ± 0.014	0.022 ± 0.018
tanh	P13+WMAP(P)	0.090 ± 0.013	–

mum found in the chain samples, not the true maximum in the parameter space. Nonetheless, the difference in likelihood of these two chain maximum models is $2\Delta \ln \mathcal{L} = 5.3$ showing that the preference is not just an artifact of parameter volume and priors.

In Fig. 9 we show that the difference in C_ℓ^{EE} is that the PC ML has a much broader reionization bump that extends to higher multipoles. This directly corresponds to the preference for high redshift ionization since the angular scale of the feature is determined by the horizon at the redshift of scattering (see [12], Fig. 2). In the tanh family of models, tight constraints at $\ell < 10$ require low power between $10 < \ell < 30$ regardless of whether the data prefers it.

We also show in Fig. 9 that the projection of the tanh ML onto the 5 PC basis accurately captures the reionization feature in the polarization spectrum with deviation of $\Delta C_\ell^{EE}/C_\ell^{EE} < 0.025$ for $\ell < 40$ which tests the completeness of the 5 PC basis. In fact most of the deviation is at $\ell \sim 10 - 20$ where the tanh model has minimal power so that the fractional accuracy still reflects a high absolute accuracy.

IV. IONIZATION HISTORY LIKELIHOOD

One of the main benefits of our PC reionization analysis is that it completely encapsulates the information from the large angle polarization measurements on the ionization history within the given redshift range. With this single analysis, it is possible to infer constraints on any given reionization model in the same range. In this section, we provide a concrete prescription for an effective likelihood function for the data given an $x_e(z)$ model.

The MCMC PC analysis of the previous section returns the posterior probability density in the space of PC amplitudes m_a given the data and flat priors on the parameters, marginalized over cosmological parameters. Invoking Bayes' theorem, we can reinterpret it as an effective likelihood of the data given m_a. To determine the likelihood given $x_e(z)$ instead, we simply need to project it onto the PC basis using Eq. (2).

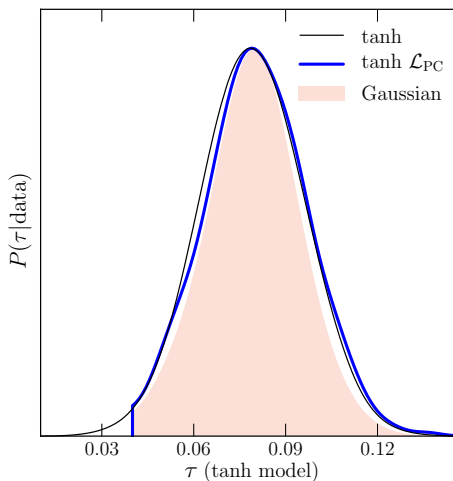In practice, the MCMC only provides a sample of the PC posterior, represented by discrete elements in the

FIG. 10. Effective likelihood analysis ($\mathcal{L}_{\mathrm{PC}}$ thick blue) on the tanh model compared with direct constraints (thin black). Even though the tanh model lives in the tails of the PC posterior, the effective likelihood models the direct constraint to a fraction of its width. Also shown is the Gaussian approximation to the effective likelihood (shaded region). The distributions are truncated at low optical depth since the fact that the universe must be reionized by $z = 6$ is enforced by the PC decomposition.

chain with parameter values $\mathbf{m}_i = \{m_1, \ldots, m_5\}$ and multiplicities w_i, whereas a given model produces a continuum of values for $\mathbf{m}$. Given the samples, we approximate the effective likelihood with a kernel density estimator of the form

$$\mathcal{L}_{\mathrm{PC}} \left(\mathrm{data}|\mathbf{m}\right) = \sum_{i=1}^{N} w_i K_f(\mathbf{m} - \mathbf{m}_i), \qquad (14)$$

where N is the total number of elements in the chain and the overall normalization is arbitrary. Here K_f is a smoothing kernel that makes the function estimate continuous at the expense of artificially broadening the distribution. We choose the shape of K_f to be a multivariate Gaussian of zero mean and covariance $f\mathbf{C}$ where $\mathbf{C}$ is the m_a covariance matrix estimated from the chain from Tab. I and Eq. (10). For a Gaussian posterior, the effect of smoothing is to increase the covariance by $1 + f$ or the errors by approximately $1 + f/2$. To minimize the amount of smoothing required to capture the behavior of models in the tail of the distribution like the standard tanh model, we oversample the posterior by running the chain past normal convergence requirements for a total of $N \approx 1.4 \times 10^6$ chain elements. In practice we choose $f = 0.14$. Note also that we employ the full chain without physicality priors since the smoothing kernel transfers information across these boundaries.

To illustrate and test this approach, we use the effective likelihood to compare constraints on the standard tanh model. In this case the model parameter is the total τ given the tanh ionization history $x_e(z; \tau)$. We can then construct the posterior probability of τ as usual via

Bayes' theorem and the effective likelihood given $\mathbf{m}(\tau)$

$$P(\tau|\mathrm{data}) \propto \mathcal{L}_{\mathrm{PC}} \left[\mathrm{data}|\mathbf{m}(\tau)\right] P(\tau). \qquad (15)$$

To match the MCMC analysis of the standard tanh model we take flat priors $P(\tau) = 1$. For the conversion between the ionization history and τ we take the cosmological parameters of the ML tanh model.

In Fig. 10, we compare the posterior probabilities from the direct MCMC analysis and the effective PC likelihood. The distributions match to much better than 1σ in their means and widths. This is a fairly stringent test on the method given that tanh models live in the tails of the PC posterior. The cutoff at low τ in the effective likelihood method simply reflects the $z_{\min}$ restriction for the PCs which assumes hydrogen ionization occurred at $z > 6$ as is observationally the case.

It is also interesting to compare these results to an even simpler effective likelihood. In Fig. 10 (shaded region), we also show the result of approximating the m_a posterior as a multivariate Gaussian with mean $\bar{m}_a$ and covariance $\mathbf{C}$

$$\mathcal{L}_{\mathrm{G}} \left(\mathrm{data}|\mathbf{m}\right) = \frac{e^{-\frac{1}{2}(\mathbf{m} - \bar{\mathbf{m}})^T \mathbf{C}^{-1}(\mathbf{m} - \bar{\mathbf{m}})}}{\sqrt{(2\pi)^5 |\mathbf{C}|}}. \qquad (16)$$

For an extremely fast but approximate effective likelihood and for models near the peak of the distribution, the Gaussian approximation may suffice.

Validated on the tanh model, our effective likelihood technique allows a rapid exploration of other models without the need for a separate MCMC analysis. To illustrate this usage, we consider the power-law (PL) models that the Planck Collaboration also analyzed for the HFI data [10]. The PL model

$$\frac{x_e(z)}{1 + f_{\mathrm{He}}} = \begin{cases} 1, & 6 < z < z_{\mathrm{end}}, \\ \left(\frac{z_{\mathrm{early}} - z}{z_{\mathrm{early}} - z_{\mathrm{end}}}\right)^{\alpha}, & z_{\mathrm{end}} \leq z \leq z_{\mathrm{early}}, \\ 0, & z_{\mathrm{early}} < z < z_{\max}, \end{cases} \qquad (17)$$

allows for an extended and asymmetric ionization history unlike the tanh model. Here $z_{\mathrm{early}} < z_{\max}$ is a parameter that truncates reionization at z_{early} but has little effect at lower redshift. In order to mimick the analysis in Ref [10] but allow for the weaker constraints from Planck 2015, we fix $z_{\mathrm{early}} = 23.1$ so that it does not significantly impact the analysis. The power law index α controls the duration of reionization Δz defined so that $x_e(z_{\mathrm{end}} + \Delta z) = 0.1(1 + f_{\mathrm{He}})$. It is therefore more convenient to parameterize the model with z_{end} and Δz. We employ the effective PC likelihood approach with flat priors for this two parameter family within their allowed ranges.

In Fig. 11, we show the 2D posteriors in these parameters. The Planck 2015 data do indeed allow an extended period of reionization if $z_{\mathrm{end}} \sim 6$ but do not particularly favor it over a prompt reionization that maintains the same total optical depth. Like the tanh model, the
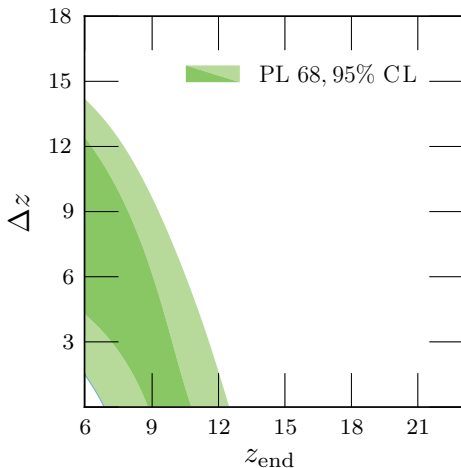
FIG. 11. Effective likelihood constraints on the two parameters of the power law (PL) model: end (z_{end}) and duration (Δz) of reionization in redshift. The PL functional form forces the two parameters to be anticorrelated given the constraint on the total optical depth and hence does not allow a separate component of high-z ionization.
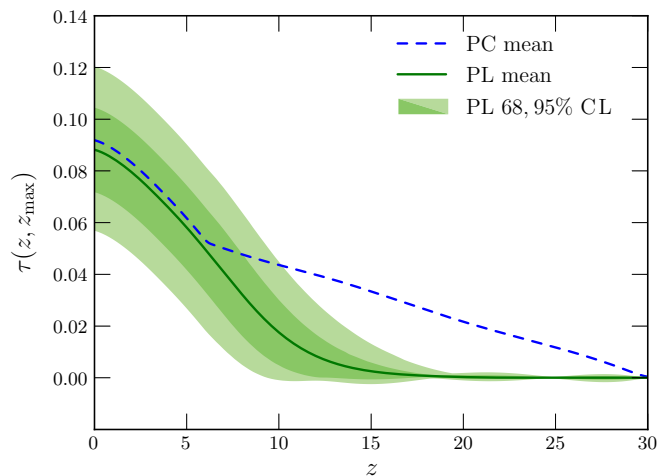

FIG. 12. Effective likelihood constraints on the cumulative optical depth $\tau(z, z_{max})$ for the PL model. Although the PL model allows an adjustable duration of reionization, its functional form does not permit the high redshift optical depth of the PC mean (blue dashed) without violating the total optical depth constraint.

PL model links the low redshift and high redshift ionization history by assumption of a functional form appropriate for models with a single phase of reionization. One should therefore not equate a constraint on the duration of reionization in the PL model with a constraint on high redshift ionization in a more general context. To see this more quantitatively, in Fig. 12 we show the cumulative optical depth $\tau(z, z_{max})$ constraints in the PL context. The PL models allow very little optical depth at $z > 15$ compared with the mean in the full PC space from Fig. 5.

The PL example also illustrates the fact that to encompass the region favored by the PC effective likeli-

hood, a reionization model would need to have an additional source of high redshift ionization that is not directly linked in functional form to its low redshift behavior. We explore such cases in a separate work [24]. In fact, we also expect the effective likelihood method to work even better for models that are favored by the data since the underlying MCMC sample better represents these models.

V. DISCUSSION

By analyzing the Planck 2015 data with an observationally complete PC basis for the ionization history, we show that it allows and even favors high redshift, $z \gtrsim 15$, optical depth at the $\sim 2\sigma$ level. The standard analysis which includes just the total optical depth and assumes a sharp step-like transition excludes this possibility by prior assumption of form rather than because it is required by the data. The same is true for power law models that additionally varies the duration of reionization.

While a 2σ result amongst the 5 PC parameters is not on its own surprising, it originates from the first and best constrained component and hence has consequences for the total optical depth. The total optical depth is important for understanding a host of other cosmological parameters from the amplitude of the current matter power spectrum σ_8 to the inferences from CMB lensing. At the very least, this analysis highlights the need for a complete treatment of CMB reionization observables to guarantee a robust interpretation of the optical depth.

This preference for extra high redshift optical depth mainly originates from the large angle polarization spectrum in the Planck 2015 data and appears related to excess power in the multipole range $10 \lesssim \ell \lesssim 20$. It is only slightly weakened by marginalizing gravitational lensing information in the temperature power spectrum, which is known to favor a higher optical depth, but more significantly changed by replacing the Planck LFI with WMAP9 polarization data.

While excess polarization power in this range favors additional sources of high redshift ionization such as population III stars or dark matter annihilation it could also indicate contamination from systematics and foregrounds. The latter have been significantly improved in the as yet proprietary Planck 2016 intermediate results. These results indicate that the low redshift end of the optical depth as tested by step-like models, or equivalently the low ℓ polarization power, is both better measured and lower than the central value in the Planck 2015 data [9]. On the other hand, these results exacerbate the tension with gravitational lensing in the shape of the temperature power spectrum which probes the total optical depth. It will be interesting to see if this complete analysis still prefers an additional high redshift component in the final Planck release.

Regardless of the outcome of resolving the mild tension between step-like reionization scenarios and the Planck

2015 data, the complete PC approach developed here is useful because with a single analysis one can infer constraints on the parameters of any reionization model within the specified redshift range, here $6 < z < 30$, but easily extensible to any desired range. What has presented an obstacle for this approach in the past is the lack of tools for converting posterior parameter constraints on the PCs to parameter constraints on models and so be able to combine them with other sources of reionization information. For example, the ionization history can also be tested in the CMB through the kinetic SZ effect from temperature fluctuations beyond the damping scale, but in a manner that is highly model dependent (e.g. [10, 25–27]).

Towards this end, we have developed and tested an effective likelihood code for inferring constraints on any given ionization history provided by a model. This approach should be especially useful in constraining models where small high redshift contributions to the optical depth need to be separated from the total.

[1] A. Mesinger, ed., *Understanding the Epoch of Cosmic Reionization: Challenges and Progress*, Astrophysics and Space Science Library, Vol. 423 (2016).

[2] W. Hu and B. Jain, Phys. Rev. **D70**, 043009 (2004), arXiv:astro-ph/0312395 [astro-ph].

[3] K. M. Smith, W. Hu, and M. Kaplinghat, Phys. Rev. **D74**, 123002 (2006), arXiv:astro-ph/0607315 [astro-ph].

[4] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, Phys. Rev. **D92**, 123535 (2015), arXiv:1509.07471 [astro-ph.CO].

[5] M. J. Mortonson and W. Hu, Phys. Rev. **D80**, 027301 (2009), arXiv:0906.3016 [astro-ph.CO].

[6] M. J. Mortonson, C. Dvorkin, H. V. Peiris, and W. Hu, Phys. Rev. **D79**, 103519 (2009), arXiv:0903.4920 [astro-ph.CO].

[7] M. J. Mortonson and W. Hu, Phys. Rev. **D77**, 043506 (2008), arXiv:0710.4162 [astro-ph].

[8] A. Kogut *et al.* (WMAP), Astrophys. J. Suppl. **148**, 161 (2003), arXiv:astro-ph/0302213 [astro-ph].

[9] N. Aghanim *et al.* (Planck), (2016), arXiv:1605.02985 [astro-ph.CO].

[10] R. Adam *et al.* (Planck), (2016), arXiv:1605.03507 [astro-ph.CO].

[11] K. Ahn, I. T. Iliev, P. R. Shapiro, G. Mellema, J. Koda, and Y. Mao, Astrophys. J. Lett. **756**, L16 (2012), arXiv:1206.5007.

[12] W. Hu and G. P. Holder, Phys. Rev. **D68**, 023001 (2003), arXiv:astro-ph/0303400 [astro-ph].

[13] A. Lewis, J. Weller, and R. Battye, Mon. Not. Roy. Astron. Soc. **373**, 561 (2006), arXiv:astro-ph/0606552 [astro-ph].

[14] M. J. Mortonson and W. Hu, Astrophys. J. **672**, 737 (2008), arXiv:0705.1132 [astro-ph].

[15] M. J. Mortonson and W. Hu, Astrophys. J. **686**, L53 (2008), arXiv:0804.2631 [astro-ph].

[16] P. A. R. Ade *et al.* (Planck), Astron. Astrophys. **571**, A22 (2014), arXiv:1303.5082 [astro-ph.CO].

[17] W.-M. Dai, Z.-K. Guo, and R.-G. Cai, Phys. Rev. **D92**, 123521 (2015), arXiv:1509.01501 [astro-ph.CO].

[18] G. D. Becker, J. S. Bolton, and A. Lidz, Publ. Astron. Soc. Austral. **32**, 45 (2015), arXiv:1510.03368 [astro-ph.CO].

[19] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].

[20] G. D. Becker, J. S. Bolton, M. G. Haehnelt, and W. L. W. Sargent, Mon. Not. Roy. Astron. Soc. **410**, 1096 (2011), arXiv:1008.2622 [astro-ph.CO].

[21] N. Aghanim *et al.* (Planck), Submitted to: Astron. Astrophys. (2015), arXiv:1507.02704 [astro-ph.CO].

[22] W. Hu, Phys. Rev. **D65**, 023003 (2002), arXiv:astro-ph/0108090 [astro-ph].

[23] P. A. R. Ade *et al.* (Planck), (2015), arXiv:1502.01589 [astro-ph.CO].

[24] V. Miranda, A. Lidz, C. H. Heinrich, and W. Hu, (2016), arXiv:1610.00691 [astro-ph.CO].

[25] M. J. Mortonson and W. Hu, Phys. Rev. **D81**, 067302 (2010), arXiv:1001.4803 [astro-ph.CO].

[26] O. Zahn *et al.*, Astrophys. J. **756**, 65 (2012), arXiv:1111.6386 [astro-ph.CO].

[27] N. Battaglia, A. Natarajan, H. Trac, R. Cen, and A. Loeb, Astrophys. J. **776**, 83 (2013), arXiv:1211.2832 [astro-ph.CO].