SEMI-BLIND EIGEN ANALYSES OF RECOMBINATION HISTORIES USING COSMIC MICROWAVE BACKGROUND DATA

M. FARHANG1,2, J. R. BOND1, AND J. CHLUBA1
1 Canadian Institute for Theoretical Astrophysics, 60 St. George, Toronto, ON M5S 3H8, Canada
2 Department of Astronomy and Astrophysics, University of Toronto, 50 St. George, Toronto, ON M5T 2P3, Canada

Received 2011 October 20; accepted 2012 April 12; published 2012 May 30

ABSTRACT

Cosmological parameter measurements from cosmic microwave background (CMB) experiments, such as Planck, ACTPol, SPTPol, and other high-resolution follow-ons, fundamentally rely on the accuracy of the assumed recombination model or one with well-prescribed uncertainties. Deviations from the standard recombination history might suggest new particle physics or modified atomic physics. Here we treat possible perturbative fluctuations in the free electron fraction, $X_e(z)$, by a semi-blind expansion in densely packed modes in redshift. From these we construct parameter eigenmodes, which we rank order so that the lowest modes provide the most power to probe $X_e(z)$ with CMB measurements. Since the eigenmodes are effectively weighted by the fiducial $X_e$ history, they are localized around the differential visibility peak, allowing for an excellent probe of hydrogen recombination but a weaker probe of the higher redshift helium recombination and the lower redshift highly neutral freezeout tail. We use an information-based criterion to truncate the mode hierarchy and show that with even a few modes the method goes a long way from the fiducial recombination model computed with Recfast, $X_{e,f}(z)$, toward the precise underlying history given by the new and improved recombination calculations of CosmoRec or HyRec, $X_{e,i}(z)$, in the hydrogen recombination regime, though not well in the helium regime. Without such a correction, the derived cosmic parameters are biased. We discuss an iterative approach for updating the eigenmodes to further hone in on $X_{e,i}(z)$ if large deviations are indeed found. We also introduce control parameters that downweight the attention on the visibility peak structure, e.g., focusing the eigenmode probes more strongly on the $X_e(z)$ freezeout tail, as would be appropriate when looking for the $X_e$ signature of annihilating or decaying elementary particles.

Key words: cosmic background radiation – methods: data analysis

Online-only material: color figures

1 INTRODUCTION

The Planck Surveyor1 is now well into its mission, observing the temperature and polarization anisotropies of the cosmic microwave background (CMB) with unprecedented accuracy (Planck HFI Core Team et al. 2011; Mennella et al. 2011). Both the Atacama Cosmology Telescope (ACT; e.g., see Hajian et al. 2011; Dunkley et al. 2011; Das et al. 2011) and the South Pole Telescope (SPT; Lueker et al. 2010; Vanderlinde et al. 2010) are pushing the frontier of TT CMB power spectra at small scales, and in the near future SPTPol4 (McMahon et al. 2009) and ACTPol5 (Niemack et al. 2010) will provide additional small-scale $E$-mode polarization data, complementing the polarization power spectra obtained with Planck and further increasing the significance of TT power spectra.

Using these data sets, cosmologists will be able to determine the key cosmological parameters with high precision (The Planck Collaboration 2006; Tauber et al. 2010), making it possible to distinguish between various models of inflation (e.g., see Komatsu et al. 2011 for recent constraints from WMAP) by measuring the precise value of the spectral index of scalar perturbations, $n_s$, and constraining its possible running, $n_{run}$, as well as the tensor-to-scalar ratio, $r$. In addition many non-standard extensions of the minimal inflationary model are under discussion, and the observability of these possibilities with Planck (The Planck Collaboration 2006) and future CMB experiments is being considered.

These encouraging observational prospects have motivated various independent groups (e.g., see Dubrovich & Grachev 2005; Chluba & Sunyaev 2006b; Kholupenko & Ivanikh 2006; Switzer & Hirata 2008; Wong & Scott 2007; Rubiño-Martín et al. 2008; Karshenboim & Ivanov 2008; Hirata 2008; Chluba & Sunyaev 2008; Jentschura 2009; Labzowsky et al. 2009; Grin & Hirata 2010; Ali-Haïmoud & Hirata 2010) to assess how uncertainties in the theoretical treatment of the cosmological recombination process could affect the science return of Planck and future CMB experiments. The precise evolution of the free electron fraction, $X_e$, with time influences the shape and position of the peak of the Thomson visibility function, which defines the last scattering surface (Sunyaev & Zeldovich 1970; Peebles & Yu 1970), and hence controls how photons and baryons decouple as electrons recombine to form neutral helium and hydrogen atoms. Consequently, the ionization history changes the acoustic oscillations in the photon–baryon fluid during recombination and therefore directly affects the CMB temperature and polarization power spectra. For the analysis of future CMB data this implies that in particular close to $z \sim 1100$ the ionization history had better be understood at the $\sim 0.1\%$ level.

Probing the ionization history in time is equivalent to probing it in space with the light cone relating the two. Thus what we try to do in this paper, namely to come up with optimized probing functions for the recombination history, is quite akin to creating probes of the spatial structure of the boundary between H II and neutral hydrogen regions. Here of course we look from neutral to ionized, the cosmological recombination problem being an inside-out H II region, except in a predominantly
electron scattering regime with a very large photon-to-baryon ratio which lowers the transition temperature between ionized and neutral.

The old recombination standard was set by Recfast (Seager et al. 1999, 2000), but its reliability for the precision cosmology was brought into question, e.g., by Seljak et al. (2003). For the standard six-parameter cosmology in particular our ability to measure the precise value of $n_s$ and the baryon content of our universe may be compromised if modifications to the recombination model of Recfast are neglected (Rubíño-Martín et al. 2010; Shaw & Chluba 2011), introducing biases of a few $\sigma$ for Planck.

Currently it appears that all important corrections to the standard recombination scenario (SRS hereafter) have been identified (see, e.g., Fendt et al. 2009; Rubíño-Martín et al. 2010, for an overview). The new recombination codes CosmoRec (Chluba & Thomas 2011) and HyRec (Ali-Haïmoud & Hirata 2011) both account for these modifications to the SRS, superseding the physical model of Recfast and allowing fast and accurate computation of the ionization history on a model-by-model basis. CosmoRec and HyRec presently agree at a level of $\lesssim 0.1\% - 0.2\%$ during hydrogen recombination, so that from standard recombination physics little room for big surprises seems to be left.

However, what if something non-standard happened? What if something was overlooked in the standard recombination scenario? From the scientific point of view the ionization history is a theoretical ingredient to the cosmological model, which usually is assumed to be precisely known and not subject to direct measurement. Clearly, it is important to estimate the possible level of uncertainty in the recombination model and to confront our understanding of the recombination problem with direct observational evidence. Here we describe how well future cosmological data alone are able to constrain possible deviations from the SRS.

In the past, several non-standard extensions of the recombination scenario have been considered. These include models of delayed recombination in which hypothetical sources of extra photons that can lead to ionizations or excitations of atoms are introduced using simple parameterizations (Peebles et al. 2000). In particular, models of decaying (e.g., see Chen & Kamionkowski 2004; Zhang et al. 2007) and annihilating particles (e.g., see Padmanabhan & Finkbeiner 2005; Zhang et al. 2006; Galli et al. 2009a, 2011; Hütsi et al. 2009, 2011; Slatyer et al. 2009) were discussed. In addition to extra photons, varying fundamental constants (e.g., see Kaplinghat et al. 1999; Scóccola et al. 2009; Galli et al. 2009b) could affect the recombination dynamics in subtle ways.

All these ideas rely on a specific model for the (physical) process under consideration, with the derived constraints depending on the chosen parameterization. This minimizes the number of additional parameters, but does not allow us to answer questions about more general perturbations around the SRS and how well they can actually be constrained.

Here we approach this problem in a different way. We introduce perturbations to the SRS over a wide range of redshifts around hydrogen ($z \sim 1100$) and helium ($z \sim 1800$) recombination, using different basis functions. We then compute the corresponding signals in the CMB power spectra and perform a principal component decomposition to obtain eigenmode functions, ordered with respect to the level at which they can be constrained by the data. We study in detail how the eigenmodes depend on the chosen parameterization for the recombination perturbations as well as the fiducial model and different experimental settings.

Our method is similar to the one used by Mortonson & Hu (2008), where the eigenmodes for different reionization scenarios ($6 \lesssim z \lesssim 30$) were constructed. However, here we explicitly construct the mode functions at redshifts $z \gtrsim 200$, with particular attention to the dependence of the eigenmodes on different assumptions. We investigate how to use our prior knowledge of possible perturbations of the ionization history to choose the parameterization which is more preferred by the data. We also carry out a careful convergence study and show the equivalence of different basis functions (e.g., triangles, Gaussian bumps, Fourier series, and Chebyshev polynomials). We particularly focus on the helium recombination problem, showing that in the absence of very tight constraints on the hydrogen recombination, we are unable to unravel well remaining uncertainties in helium recombination with CMB data. Similarly, small changes in the freezeout tail of recombination are only weakly constrained, if possible ambiguities during hydrogen recombination are included.

Details of the general methodology to construct the eigenmodes for perturbations to ionization history are given in Section 2 and Appendices A–C. In Section 3 we compute different eigenmodes over a rather wide redshift range ($z \in [200, 3000]$) and investigate their properties. In Section 3.6 we develop a criterion which allows us to truncate the hierarchy of the eigenmodes based on their information content. In Section 4 the modes are applied to two specific examples of ionization scenarios, illustrating how the method should be used with real CMB data. At the end of that section, we also discuss how the approach should be iterated if hints toward a considerable difference between the assumed and true model of recombination are indicated by the data. We close the paper with a brief discussion.

2. METHODOLOGY

In this section we introduce the approach and parameterization used to construct the principal components, or the eigenmodes, which are later used to describe possible corrections to the recombination scenario. Our method is mainly driven by the assumption of small relative perturbations around the fiducial model computed with the Recfast code (see, e.g., Seager et al. 1999; see Wong et al. 2008 for recent updates). As an example we have in mind the recombination corrections obtained with refined recombination models (Chluba & Thomas 2011; Ali-Haimoud & Hirata 2011). However, we also briefly discuss the possibility to constrain significant changes in the freezeout tail of recombination and modes that mainly focus on helium recombination.

Throughout this paper the cosmic parameters, referred to as the standard (cosmological) parameters, are ($\Omega_b h^2$, $\Omega_m h^2$, $H_0$, $\tau$, $n_s$, and $A_s$) as measured by WMAP7, unless stated otherwise. In several cases we also vary Yp as a seventh parameter. Lensing is included in all simulations if not explicitly stated otherwise.

2.1. The Standard Recombination Scenario

The cosmological recombination history is one of the major theoretical inputs for computations of the CMB anisotropies. Consequently, high-precision unbiased cosmic parameter measurements from current and future CMB experiments require a sufficiently accurate model for hydrogen and helium recombination.

http://lambda.gsfc.nasa.gov/product/map/dr4/params/lcdm_sz_lens_wmap7.cfm
The ionization fraction for the SRS is shown in the left panel of Figure 1. It was calculated using Recfast v1.4.2, which accounts for some of the modification to helium recombination (Kholupenko et al. 2007; Switzer & Hirata 2008; Rubíno-Martín et al. 2008; Chluba & Sunyaev 2010) using fudge parameters, but neglects detailed radiative transfer corrections (see Chluba & Thomas 2011; Ali-Ha¨ımoud & Hirata 2011, and references therein) around \( z \sim 1100 \). The solid curve corresponds to an ionization fraction with the measured temperature of the CMB radiation, \( T_{\text{CMB}} \sim 2.726 \) K (Fixsen 2009). For comparison and to illustrate the temperature dependence of the ionization history, the ionization fraction corresponding to \( T_{\text{CMB}} = 3 \) K is also plotted (dashed curve). A larger value of \( T_{\text{CMB}} \) means more photons in the Wien tail of the CMB blackbody, so that the matter is kept ionized even lower redshift.

On the right the corresponding differential visibility functions (or visibility functions for short) are plotted:

\[
g(z) \equiv \frac{d e^{-\tau(z)}}{d \eta},
\]

where \( \eta \) is the conformal time and \( \tau \) is the Thomson scattering optical depth from redshift \( z \) to now.

The visibility function describes the probability that a photon we observe today last scattered off free electrons at a certain position along the line of sight. The CMB anisotropies formed mainly during the epoch of hydrogen recombination defined by the peak of the visibility function located at redshift \( z \sim 1100 \). They are thus most sensitive to changes around the maximum of visibility. For example, an increase in the width of the visibility bump corresponds to a more extended or slower recombination process, leading to more Thomson scatterings of photons off free electrons. These scatterings lead to the cancellation of the CMB anisotropies along the line of sight on scales comparable to and smaller than the recombination width, while enhancing the polarization signal on larger scales. The location of the maximum of the visibility function for an assumed cosmological model, on the other hand, determines the distance to the last scattering surface. This in turn affects the positions of the peaks of the CMB power spectra. Similarly, any change in the ionization history, through affecting the visibility, leads to (possibly measurable) changes in the CMB power spectra.

As the right panel of Figure 1 indicates, at high redshifts \( z \gtrsim 1400 \) the visibility function falls off very quickly. At those times the number of free electrons is still so large that scatterings occur very frequently and the mean free path is very short. Consequently, the part of the ionization history which is connected to helium recombination mainly affects the damping tail of the CMB anisotropies, but even there the effect is rather moderate, as in the redshift range \( 1400 \lesssim z \lesssim 3000 \) helium can at most alter the number of electrons by \( \sim 8\% \).

### 2.2. Choice of Perturbation Parameterization

There are different ways to parameterize possible deviations from the assumed fiducial ionization history in a (semi-)model-independent way. For example, to study how well the low-redshift ionization history (6 \( \lesssim z \lesssim 30 \)) can be constrained by future CMB data, Hu & Holder (2003) and Mortonson & Hu (2008) introduced changes in the ionization fraction in different redshift bins with \( \delta X_e(z) = \) constant to parameterize the uncertainties. This is a valid choice for the low-redshift region because our ignorance of the underlying model of reionization does not suggest any preferred non-uniform weighting of the perturbations at different redshifts. In this regime \( \delta X_e(z) \) probes the ionization fraction itself and not perturbations guided by a fiducial model. The results from this choice of parameterization are shown to be fiducial model independent which is expected due to the weak signal from the reionization process.

In contrast to this, at high redshifts (\( z \sim 1100 \)) there is strong theoretical support for the exhaustively studied model of recombination in the realm of standard atomic physics and radiative processes. Also, the current generation of CMB data is sensitive to changes in \( X_e \) at the level of a few percent. Therefore the main assumption in this paper is that the fiducial model for the ionization history, \( X_e^{\text{fid}}(z) \), is close to the true underlying history, \( X_e(z) \), which we are looking for. We call this method semi-blind, emphasizing our belief in the SRS as the framework of recombination, with the search for deviations being limited to small perturbations around this reference model. The goal is to detect or place upper limits on possible small deviations. Clearly, if data point toward significant deviations from the SRS, an iterative approach should be adopted, as discussed in Section 4.4.

\( ^7 \) The recombination of doubly ionized helium ends around redshift \( z \sim 5000 \).
With small deviations in mind we can write
\[ X_e(z) = X_e^{\text{fid}}(z) + \delta X_e(z), \]
with \(|\delta X_e(z)|/X_e^{\text{fid}} \ll 1\). A natural parameter to describe the perturbation is then the relative deviation in the ionization fraction:
\[ \delta u(z) \equiv \frac{\delta X_e(z)}{X_e^{\text{fid}}(z)} \quad \text{with} \quad |\delta u(z)| \ll 1. \]  
(1)

This parameterization has the advantage of always satisfying the necessary condition \( X_e \geq 0 \). It is also straightforward to fulfill \( X_e \leq X_{e,\text{max}} \) in the simulations, where \( X_{e,\text{max}} \) is determined by \( Y_p \), the primordial helium mass abundance, through \( X_{e,\text{max}} \approx 1 + Y_p/2(1 - Y_p) \). The parameterization in Equation (1) weights possible perturbations at different redshifts by the fiducial ionization fraction. This implies that for \( \delta u(z) = \) constant the absolute change in the ionization fraction \( |\delta X_e| \) is downweighted in the freezeout tail of \( X_e (z \lesssim 800; \text{see Figure 1}), \) compared to perturbations around maximum visibility \((z \sim 1100)\) where \( X_e \) rapidly approaches unity. Throughout this paper, \( \delta u(z) \) as defined in Equation (1) is our main choice of parameterization.

A more general parameter which includes the above parameterization as a special case is given by
\[ \delta u(z) \equiv \frac{\delta X_e(z)}{[X_e^{\text{fid}}(z) + \sigma(z)]}, \]  
(2)

where \( \sigma(z) \geq 0 \) can be a constant or otherwise convenient function of redshift allowing to focus on different redshift ranges of interest. In particular, when considering possible modifications to the ionization history introduced by energy injection, e.g., because of annihilating dark matter (DM), or decaying relic particles (Chen & Kamionkowski 2004; Padmanabh & Finkbeiner 2005; Zhang et al. 2006, 2007; Hütsi et al. 2009, 2011; Slatyer et al. 2009; Galli et al. 2009a, 2011), where the freezeout tail of recombination is disturbed, a value of \( \sigma \gg X_e^{\text{fid}} \) might be a good choice, giving higher weight to the perturbations in the lower redshift tail (see Section 3.1 and Figure 12). In the limit of a high value of \( \sigma \) relative to the fiducial \( X_e \) the parameters approach \( \delta u(z) = \delta X_e(z) \) which uniformly weights perturbations at different redshifts. This, as already discussed, is a good choice for regions where there is no strong a priori belief in the underlying model or if the redshifts of interest have comparatively low \( X_e \) where \( \delta u(z) \) with \( \sigma = 0 \) does not lead to strong enough signals to probe. In principle, a conveniently chosen redshift-dependent \( \sigma(z) \) is a tool to effectively incorporate our prior knowledge of the ionization history in the parameterization of its perturbations. For example, with \( \delta u(z) \) defined by Equation (2) one can smoothly interpolate between relative and absolute perturbations to \( X_e \), at high and low redshifts, respectively. Also it is clear that one can focus on different parts of the recombination history by limiting the redshift range over which the eigenmodes are constructed, e.g., just on reionization \((0 \lesssim z \lesssim 30)\) or helium recombination \((1400 \lesssim z \lesssim 3000)\).

2.2.1. Alternative Parameterizations

We comment that instead of directly perturbing the ionization fraction, as we chose here, it is plausible to parameterize possible changes in the physical sources of perturbation to the ionization history, such as energy injection in the medium which leads to excitation or ionization of atoms, or the Lyα escape probability during recombination (see the photons). For example Mitra et al. (2011) chose the number of photons in the intergalactic medium per baryon in collapsed objects as the parameter to study the low-redshift ionization history. Alternatively, one could modify the fudge factors or functions in Recfast, or alter the expansion rate given by the Hubble factor, \( H(z) \).

Each of these possibilities implies different priors on the regions that can be altered and, e.g., in the case of \( H(z) \), other aspects of the cosmological model are also affected. They also cover, in general, only a limited class of changes to the recombination history. When interested in perturbations to the ionization history, \( X_e \) is the physical quantity which, via the visibility function and the optical depth, most directly enters the Boltzmann equations describing the evolution of radiation anisotropies routinely solved using the Boltzmann codes such as CAMB (Lewis et al. 2000) or CMBfast (Seljak & Zaldarriaga 1996).

The ionization fraction has the additional advantage, over the visibility and the optical depth, of being straightforward to limit to physically allowed values. The nearly direct mathematical encounter of \( X_e \) with CMB anisotropies guarantees that any perturbation in the plasma that would lead to changes in the ionization anisotropies should go through and thus be reflected in \( X_e \). Therefore, the relative changes in \( X_e \) constitute our preferred physical parameters.

We close by mentioning that it is also theoretically possible to consider different variables for time such as (conformal) time, optical depth, and scale factor. However, for our purpose we choose to work with redshift to describe temporal dependence. In principle, different parameterizations, if they cover the same range of physical perturbations, can be transformed to one another with the proper change of the a priori distribution of parameters. Here, in the absence of physically motivated constraints, a uniform prior is assumed for perturbations at different redshifts regardless of parameterization (here, e.g., for various values of \( \sigma \) in Equation (2)). If the perturbation is strongly constrained by data, the choice of the prior does not play a major role.

2.3. Basis Functions and Their Different Characteristics

Having chosen the parameterization, we now expand the perturbations in a discrete set of mode (or basis) functions, \( \psi_i(z) \):
\[ \delta u(z) = \sum_{i=1}^{N} y_i \psi_i(z) + r(z) \quad \text{with} \quad z_{\text{min}} \leq z \leq z_{\text{max}} \]  
(3)

and \( \delta u(z) = 0 \) elsewhere. Here \( r(z) \) is the residual and \( y_i \) are the parameters defining the strength of the mode \( \psi_i(z) \). Often we take \( \psi_i(z) \) to be localized in \( z \) about a knot value \( z_i \), but this is not necessary. We can, for example, choose the \( \psi_i(z) \) to form a complete orthonormal set in which case \( N \to \infty \) and the residual \( r \) approaches zero. Below we discuss different possibilities for the mode functions. We modified the publicly available code CAMB to simulate CMB power spectra for a more general recombination scenario that includes perturbations on top of the SRS. Introducing narrow features into the ionization history also required an increase in the redshift sampling of \( X_e \). We checked the numerical convergence and stability of the results by using high accuracy settings.

Localised basis functions. We first investigated three sets of localized basis functions: Gaussian and triangular bumps, which can be considered as approximations to the Dirac \( \delta \)-function, and \( M_4 \) splines, a commonly used kernel in smoothed particle
Localized perturbations in the $X_e$ history, in the form of $M_4$ splines (left) and the derivatives of the $C_l$ with respect to the amplitude of each perturbation ($TT$ power spectrum in the center, $EE$ on the right).

(A color version of this figure is available in the online journal.)

Figure 2. Similar to Figure 2 but for non-localized perturbations in the form of Chebyshev polynomials.

(A color version of this figure is available in the online journal.)

hydrodynamics (SPH). For a detailed description of these functions and how their properties compare, see Appendix A. As examples for localized perturbations, Figure 2 shows three perturbation functions $\delta u(z) = \delta \ln(X_e)$ using $M_4$ splines (left panel) and the corresponding $C_l$ response in $TT$ and $EE$. The perturbations are located at different redshifts and have equal widths. We see that the amplitude of the response typically increases at smaller scales indicating a change in the duration of the recombination epoch (i.e., the effective width of the visibility function). The $C_l$ response also has an oscillatory component similar to a change of the position of the visibility peak. These oscillations are most noticeable for the perturbations close to the visibility peak ($z \sim 1100$).

Non-localized basis functions. We also expanded the perturbations in terms of two non-localized basis functions, namely, a Fourier series and Chebyshev polynomials. For more details see Appendix B. The non-local basis functions are very different in nature from the localized ones discussed above. Therefore, the response of the observables (here the $C_l$'s) to the perturbation $\delta u(z)$ in the form of these functions is also expected to be rather different. Figure 3 shows the $C_l$ responses when perturbing the ionization history using Chebyshev polynomials with different frequencies. We see that perturbations with low frequencies, covering a large redshift range, lead to $C_l$ responses with much larger mean amplitudes when compared to the perturbations in the form of local bumps (Figure 2). However, as the frequency of the oscillations of the basis function increases, the response becomes weaker and its oscillations damp away. That is because neighboring oscillations lead to similar responses in the $C_l$'s with opposite signs and can partially cancel each other. In other words, the CMB power spectra are less sensitive to high-frequency perturbations in the ionization history.

In principle, in the limit of large mode number, all bases work well (see Section 3.1). However, we found that for the recombination history, although non-localized basis sets have their virtues, the $z$-localized bases are better, especially if we are trying to describe narrow features in redshift. We return to this point in Section 3.1.

2.4. Constructing the Eigenmodes

So far we introduced our choice of parameterization for characterizing possible perturbations to $X_e$ and illustrated how different set of basis functions affect the CMB power spectra. In principle, all the amplitudes $y_i$ defined in Equation (3) are needed for a (nearly) complete reconstruction of a general perturbation $\delta u(z)$. However, in practice data cannot constrain the perturbations in detail in many cases. As shown in Figure 3, very-high-frequency perturbations are expected to lead to much smaller signals.

To avoid dealing with many correlated (and possibly weakly constrained) parameters (i.e., the $y_i$'s), we construct a set of their linear combinations which are uncorrelated with each other and only keep those combinations that are most constrained by data. This procedure provides a hierarchy of mode functions and their corresponding signals in the CMB temperature and polarization power spectra. Exclusion of the weakly constrained eigenmodes
does not affect the measurement of the rest of parameters since the eigenmodes describing the recombination perturbations are by construction uncorrelated. The standard well-developed procedure of using an orthogonal transformation to replace the parameters of the problem with a set of uncorrelated variables is called principal component analysis or PCA for short. The parameter eigenmodes were used for CMB in Bond (1996) and subsequently in Bond et al. (1997) and many subsequent papers.

In the process of eigenmode construction for the perturbations, assuming the standard parameters are known with high precision (see the discussion in Section 3.3), one only needs to deal with the perturbations in a fixed background cosmology. We call these modes $X_e$ eigenmodes, or eXeMs for short. However, for cases where the cosmic parameters are also measured simultaneously with the perturbations, these eXeMs are no longer the optimal perturbation patterns. They do not stay uncorrelated and the uncertainty in their measurement increases, as we see in Section 4.1. In such cases, the effect of the correlation of the cosmic parameters with the perturbations needs to be taken into account while constructing the eigenmodes. In other words, the modes should be marginalized over the standard cosmic parameters. We call these eigenmodes the extended $X_e$ eigenmodes or eXeMs. The eXeMs stay uncorrelated to each other (but not necessarily to the standard parameters) even in the presence of varying background cosmology. A detailed discussion of the PCA for both fixed and varying background cosmology is presented in Appendix C.

In general, the PCA needs to be applied to the whole ionization history simultaneously (as well as the standard cosmic parameters), since we do not know a priori how the ambiguity in one epoch affects the measurements of perturbations in other epochs. However, if it turns out that a particular period of $X_e$ history could be relatively well constrained, e.g., by other cosmological probes, one could leave that epoch out of perturbations. Moreover, choosing a suitable parameterization, potentially changing over time to properly take into account the different physics at different epochs, is a necessary but not straightforward task. In this work, the focus is on the epoch of recombination since that is where the main CMB signal is coming from. A more complete analysis for the whole parameter space corresponding to $X_e$ is not perturbed significantly, especially for lower values of $\sigma$. We call these modes the most correlated Xe eigenmodes, or XeMs for short. We propose an information-based criterion for truncating the eigenmode hierarchy to be used in the data analysis. Finally, in two examples, we show how these eigenmodes help reconstruct some physically motivated perturbations.

3. PERTURBATION EIGENMODES
FOR RECOMBINATION

In this section, we follow the procedure of Section 2.4 (and Appendix C) to find the eigenmodes for perturbations in the ionization fraction at high redshifts. We choose the redshift range of [200, 3000] which covers hydrogen and singly ionized helium recombination ($z \sim 1100$ and $z \sim 1800$, respectively) as well as part of the dark ages while leaving reionization ($z \lesssim 30$) unaltered. We assume that the fiducial recombination history is given by the SRS, as explained in Section 2.1, unless otherwise stated.

In the following (and in Appendix D) we compare the eigenmodes generated by using various bases and several different parameterizations and study some of the aspects associated with them, such as their convergence and fiducial model dependence. Special attention is given to perturbation to helium recombination. We also study how including the standard cosmic parameters in the analysis changes the eigenmodes.

3.1. XeM Construction

In this section, we use the perturbation parameterization $\delta u(z) = \delta \ln(X_e + \sigma)$ with various values of $\sigma$, including $\sigma = 0$. For each $\sigma$ we calculate the $N \times N$ Fisher information matrix as explained in Appendix C where the $N$ parameters are the amplitudes of the perturbations in the form of the basis functions (e.g., Gaussians), i.e., $y_i$ introduced in Equation (3). The standard cosmic parameters are fixed to their fiducial values. For the data we simulate the $TT$, $TE$, and $EE$ spectra up to $\ell = 3500$ for a full-sky, cosmic-variance-limited (hereafter CVL) CMB experiment, unless otherwise stated. We then construct the Fisher matrix (Equation (C1)) and from it the $N$ XeMs (Equation (C3)). The first six XeMs for $N = 160$ and for four values of $\sigma$ are shown in Figure 4. The first row in Table 1 shows the forecasted errors of these XeMs for $\sigma = 0$, obtained from the eigenvalues of the Fisher matrix. Note that including standard parameters in the analysis, e.g., in Markov Chain Monte Carlo (MCMC) simulations, increases the error bars, as we discuss in Section 4.

We see that the first six XeMs—which are the most constrained modes—all have the strongest variations close to the maximum of Thomson visibility function and the freezeout tail is not perturbed significantly, especially for lower values of $\sigma$. As $\sigma$ increases, however, the amplitude of the XeMs in the freezeout tail increases. This is because choosing $\sigma > 0$ results in overweighting the signal from perturbations at low $z$ (with

| XeM | 1   | 2   | 3   | 4   | 5   | 6   |
|-----|-----|-----|-----|-----|-----|-----|
| CVL ($\ell_{\text{max}} = 3500$) | 0.003 | 0.009 | 0.013 | 0.016 | 0.022 | 0.047 |
| CVL ($\ell_{\text{max}} = 2000$) | 0.011 | 0.019 | 0.024 | 0.041 | 0.094 | 0.190 |
| CVL ($\ell_{\text{max}} = 3500$, T only) | 0.004 | 0.021 | 0.064 | 0.103 | 0.208 | 0.275 |
| Planck-ACPol ($\ell_{\text{max}} = 3500$) | 0.015 | 0.047 | 0.068 | 0.13 | 0.22 | 0.31 |
The oscillations around helium recombination ($z \sim 1800$) also have much smaller amplitude than those at $z \sim 1100$. This is expected since the CMB anisotropies are most sensitive to perturbations during maximum visibility and features at low and high redshifts are not weighing as much in the CMB power spectra, if uncertainties close to $z \sim 1100$ are admitted. This in turn implies that only once the ionization history during hydrogen recombination is known well can small modifications at higher redshift or in the freezeout tail be constrained.

In Figure 4, we can observe another aspect of the XeMs: the larger the expected error bar the more high frequency oscillations the modes have and the further away from the visibility peak they probe. This is again understandable, since neighboring ups and downs in the mode functions lead to partial cancellation of the effect on the $C_\ell$. Once several oscillations are occurring close to $z \sim 1100$, signals produced farther away from maximum visibility can start competing with those from $z \sim 1100$, and hence become constrainable by the data.

One also expects the XeMs to be independent of the choice of the basis functions. We demonstrate this by trying the five different sets of basis functions of Section 2.3 (see also Appendices A and B): Chebyshev polynomials and Fourier series as orthogonal non-local functions of redshift, and $M_4$ splines, triangular and Gaussian bumps as localized basis functions. We find that the first few XeMs are practically the same independent of the chosen expansion basis (three sample XeMs are shown in Figure 5), although individual perturbations in different bases lead to totally different $C_\ell$ responses (cf. Figures 2 and 3).

The eigenmodes are also converged and do not change by including a larger number of parameters. We tested this by trying $N = 40, 80, 160$, and 320 in different bases and found that by $N = 160$ the first few modes are converged (cf. Figure 6). For the case of $M_4$ spline functions the robustness of the results should also be checked against increasing the width of the kernel. By comparing the (first six) XeMs with $h = 1.5\delta z$ (as defined in Appendix A) to those with $h = 3\delta z$, we conclude that the modes have already converged for $h = 1.5\delta z$ which we adopt for the rest of this paper.

However, as we go to modes with higher uncertainty (not shown and used here), the XeMs from different bases start to slightly differ from each other. A larger number of basis functions are required to make these higher XeMs agree as well. Moreover, we found that the higher (poorly constrained) XeMs, in particular from the extended basis expansions such as Fourier, become dominated by numerical noise. The reason is that for weakly constrained modes where the higher frequencies start to play a more important role, the impact of adjacent ups and downs from the high-frequency perturbations (e.g., sine functions) may not be resolvable well in the $C_\ell$, resulting in their net effect being dominated by numerical noise. For the localized basis functions, as long as the individual bumps are numerically resolvable, we do not find this issue because each perturbation has just one bump with no destructive neighbor.

**Figure 4.** Six most constrained XeMs with $\delta u(z) = \delta \ln(X_e + \sigma)$ as the parameter, having four different values for $\sigma$, and with Gaussian bumps as the basis functions. The maximum and width (at 68% and 95% levels) of the Thomson visibility function have been marked in all figures.

(A color version of this figure is available in the online journal.)
3.2. eXeM Construction

The XeMs discussed so far were constructed with non-varying standard parameters and therefore can be considered as the limiting case of zero errors on the standard parameters. However, as mentioned in Section 2.4, the eigenmodes become correlated when they are simultaneously being measured with the standard parameters due to their degeneracy with standard parameters. The strength of the impact of these correlations on the XeM estimation depends on the (prior) constraints on the standard parameters. It is therefore worthwhile to see how the modes and their rank ordering change if the standard parameters are allowed to vary as well.

Figure 7 illustrates the first three eXeMs constructed after marginalization over the main six and seven (including \(Y_p\)) standard parameters. For all cases considered in this section we set \(\sigma = 0\). The eXeMs have stronger high-redshift features compared to the XeMs. This implies that the degeneracy between the standard parameters and some features in the perturbations of the ionization fraction have pushed back some patterns of high significance to lower levels, opening up the room for some high redshift or higher frequency patterns which only had the chance to show up at lower significance XeMs.

The modes in the two cases shown (i.e., marginalized over six and seven standard parameters) differ only slightly. That is because \(Y_p\) is rather weakly constrained using CMB data alone and in the presence of other standard parameters its role in shaping the eigenmodes is only secondary. If, on the other hand, we hold the six standard parameters fixed and only let \(Y_p\) vary, the eigenmodes are more significantly affected (see Figure 8). The reason is that \(Y_p\) is comparable in significance to small changes in the ionization fraction and marginalizing over it, without the dominance of the standard parameters, leads to marked changes in the eigenmodes. The forecasted errors on the first six eXeMs with and without \(Y_p\) included are compared in Table 2. We see that the errors are mostly the same, again implying the subdominant role of \(Y_p\). In terms of the errors the most affected modes are eXeM 2 and 6.

It is instructive to see how the eXeMs can be constructed from the XeMs. Table 3 shows the coefficients of the projection of the first six eXeMs on the best eight XeMs. Note that the most constrained eXeMs have their strongest projections along these first few XeMs and the contribution from all other modes, i.e., higher than the eighth mode, is at most about a percent for these first six eXeMs. This means that allowing the

For more precise computations of the higher XeMs improvements of the numerical treatment in CAMB would become necessary. We tried several obvious modifications, as well as different settings for the accuracy level, but were unable to stabilize the results for very high frequency modes. However, since in the analysis we are hardly using more than a few XeMs, for the purpose of this work this was sufficient.

In Appendix D, we also studied how the eigenmodes respond to changes in the fiducial model and simulated CMB data set used for their construction.
standard parameters to vary mixes and rearranges the first few XeMs with negligible leakage from higher neglected XeMs. This implies that using similar number of XeMs and eXeMs in an analysis of possible recombination perturbations should give similar results for the reconstructed modification in the ionization history, at least for the CVL case where relatively large number of eigenmodes are included. However, it also turns out that the eXeMs perform better than the XeMs in the simulated analysis of Planck data, where only 1–3 modes seem to be constrainable. The main advantage of the eXeMs is that one can obtain more realistic estimates for the error bars directly after the construction of these modes.

3.3. From eXeMs to XeMs

The two sets of ionization perturbation eigenmodes introduced and constructed so far, i.e., the XeMs and the eXeMs, allow us to best describe and measure the uncertainties in the ionization fraction in the two extreme ends of our knowledge of the standard cosmic parameters. The eXeMs present a case where the tightest constraints on the standard (six) parameters are from CMB data alone. Therefore, a simultaneous measurement of the standard parameters and the uncertainties in the ionization fraction, using the CMB data set at hand, is required. The construction of the XeMs, on the other hand, assumes that the standard cosmic parameters are measured with high accuracy from other cosmological probes and the CMB data are

Figure 7. Three most constrained eXeMs. The solid blue lines correspond to modes constructed after marginalization over six standard parameters while for the dashed red curves \(Y_p\) is marginalized over in addition.

Figure 8. First three eXeMs with only \(Y_p\) being marginalized over (dashed blue curves). For comparison, the first three XeMs are also plotted (solid black curves).

Table 2

| eXeM          | 1   | 2   | 3   | 4   | 5   | 6   |
|---------------|-----|-----|-----|-----|-----|-----|
| CVL, marg: six std | 0.011 | 0.012 | 0.029 | 0.052 | 0.059 | 0.064 |
| CVL, marg: six std + \(Y_p\) | 0.011 | 0.027 | 0.029 | 0.052 | 0.059 | 0.071 |
| Planck-ACTPol, marg: six std | 0.058 | 0.074 | 0.189 | 0.308 | 0.439 | 0.532 |
| CVL, marg: \(n_s, A_s\) | 0.009 | 0.011 | 0.016 | 0.018 | 0.033 | 0.059 |

Note. In all cases, \(\ell_{\text{max}} = 3500\).
only used for the direct measurement of the ionization history. In other words, the XeMs, by ignoring the uncertainties in the standard parameters, extract the maximal amount of information that the CMB data would ever have to offer about the ionization fraction.

Between these two limiting cases, there is a gray region where, depending on the data set at hand, tight priors from non-CMB surveys can be imposed on some of the standard parameters while the rest are marginalized over when constructing the eigenmodes. For example, if all standard parameters, but the inflationary ones $A_s$ and $n_s$, are measured to very high precision by other probes, such as large-scale structure, baryonic oscillation, lensing, and supernova surveys (e.g., LSST,9 Pan-STARRS,10 BigBOSS, WFIRST,11 EUCLID12), the corresponding eigenmodes would be constructed after marginalization only over these inflationary parameters.

Figure 9 compares the first three XeMs with eigenmodes only marginalized over $A_s$ and $n_s$. The forecasted errors on these eigenmodes (from the Fisher analysis) are presented in the last row of Table 2. Not surprisingly, these modes have smaller errors compared to the eXeMs which have been made after marginalization over six standard parameters, and have larger errors compared to XeMs (with no standard parameter varying). These modes and similar ones after marginalization over different sets of standard parameters smoothly bridge the gap between the XeMs and the eXeMs. Depending on the data sets available at the time of real data analysis, the proper eigenmodes marginalized over the appropriate standard parameters must be constructed. With the current (and very near future) surveys, the most realistic choice is the eXeMs, constructed according to the experiment under consideration, which should be quite similar to the Planck-ACTPol-like case studied here.

3.4. Perturbations to Helium Recombination

As mentioned early in this section, the redshift range chosen in our analysis of perturbations to ionization fraction includes the recombination of singly ionized helium. Some of the most constrained XeMs we found also extend up to $z \sim 1600$. These, therefore, imply some impact from the helium recombination epoch on the XeMs.

One way to confirm this statement is to limit the redshift range of perturbations to mainly include singly ionized helium recombination, e.g., [1500, 3000], while the total $X_e$ (from both hydrogen and helium) is perturbed. We observe that the XeMs constructed this way have comparably large values at the lower redshift boundary ($z = 1500$) and would steeply go to zero if enforced by the imposed boundary conditions, e.g., by prior knowledge that only this specific redshift range of helium recombination is uncertain and the perturbations outside this range are enforced to be zero. This indicates that despite being restricted to the helium recombination epoch, the XeMs are still most sensitive to changes in the signal from the hydrogen recombination and changes in $X_e$ due to helium recombination are hardly constrainable, unless a properly chosen non-uniform prior on $\delta X_e$ is imposed.

As already emphasized, the parameters $\delta u(z)$ only characterize relative changes in $X_e$, and the full description of the ionization fraction depends also on the standard cosmic parameters as well as the relevant theoretical assumptions about the physics of recombination. Among the standard parameters, $Y_p$ has a distinct role in describing an aspect of the ionization fraction complementary to $\delta u(z)$ by determining the maximum total number of electrons available at each redshift: $N_{e,\text{max}} = N_{e,\text{max}}^H + N_{e,\text{max}}^\text{He} \approx (1 - Y_p/2)N_b$, where $N_b$ is the baryon number density. Therefore, although $Y_p$, in the first instance, requires that it be marginalized over when constructing the XeMs, due to its intimate relation with the ionization fraction it is also legitimate to treat recombination perturbations and the maximum number of electrons available at each redshift separately.

In Section 4.1 we use MCMC to measure constraints on $Y_p$ alongside the six standard parameters and the first few XeMs using (simulated) CMB data. Also in the next section, we explore in more detail how the eigenmodes change if they are marginalized over $Y_p$.

3.5. Impact of the Eigenmodes on Differential Visibility and CMB Power Spectra

It is worthwhile to see how the eigenmodes affect the visibility function and the CMB power spectra. The left panel of Figure 10
Figure 10. The $\Delta \text{vis} = (\text{vis} - \text{vis}_{\text{fid}})$, normalized to the maximum of the fiducial visibility, (left) and the relative changes in the $TT$ and $EE$ power spectra (middle and right) for the six most constrainable XeMs.

(A color version of this figure is available in the online journal.)

shows the change in the visibility function (normalized to the maximum of the fiducial visibility) for the first six XeMs. The central and right panels illustrate the relative changes in the CMB temperature and $E$-mode polarization power spectra due to these first six XeMs. The amplitudes of XeMs are chosen proportional to their corresponding $1 \sigma$ values. Thus, we expect the perturbations to lead to comparable changes in the $C_\ell$. It is remarkable that such tiny relative changes in $X_e$ (and correspondingly in the visibility) lead to potentially measurable effects ($\sim$-tenth of a percent) in the CMB power spectra. This confirms the high sensitivity of the $C_\ell$ to tiny changes in the visibility.

From Figure 10 we see that the most constrained mode, XeM 1, has an effect on the CMB power spectra consistent with changes in the width of the visibility function and a slight shift of its peak position. For this XeM, due to its narrower visibility width compared to the fiducial model, the high $\ell$ damping in the temperature and polarization anisotropies is smaller. At the same time, because of fewer scattering opportunities for the photons, this mode leads to less polarization (negative $\delta C_\ell^T$ at low $\ell$'s). Higher XeMs lead to less trivial changes in the width and position of the visibility function. However, the mainly oscillatory impact of these modes on the $C_\ell$ suggests an effective shift in the position of the visibility function. But, e.g., for mode XeM 4 the tail of the power spectrum is also less strongly damped, corresponding to a change in the effective width of the visibility.

3.6. Criteria for Truncating the Eigenmode Hierarchy

For the full reconstruction of perturbations to the ionization fraction, all eigenmodes are in principle needed since they form a complete basis set. In practice, sequentially adding modes rank ordered in the (possibly renormalized) eigenvalues $f_i^{-1} = \sigma_i^2$ of $F^{-1}$ (or the eigenvalues of $(F^{-1})_{pp}$ in the more general case where the background cosmology also changes) from low to high gives a rapidly diminishing return once one goes beyond a dozen or so. And often we can learn much from using just the first few. As more modes are added, the width covered by the allowed $X_e$ trajectories increases, as Figures 19–23 in Section 4.3 show. The errors in those standard cosmic parameters which are correlated with the $X_e$ eigen parameters also increase. On both counts, it behooves us to develop criteria for selecting which modes to keep, bearing in mind Occam’s Razor for minimizing the number of new parameters to be added. Thus, we show in Section 4 what happens when one mode, a few modes, and a handful of modes are added. To be more quantitative, we explore an additional criterion based on not allowing the Shannon entropy to increase too much as the next eigenmodes in the hierarchy are added.

The information action is defined in terms of the a posteriori probability of the variables $p_i$ and the evidence $\mathcal{E}$ as $S_{ii}(\mathbf{q}) \equiv \ln p_i^{-1} - \ln \mathcal{E}$. As explained in Section 2.4, the a posteriori probability $p_i \equiv p(\mathbf{q}|d, T)$ of variables $\mathbf{q} = (q_1, \ldots, q_N)$ given the theory space $T$ and the data sets $d$ is related to the a priori...
probability \( p_i \equiv p(q_i|T) \), the likelihood \( L(q,d,T) \equiv p(d|q,T) \), and the evidence \( E \equiv p(d|T) \) through Bayes' theorem: \( p_i = L(q,d,T) p_i / E \). The information action can then be written in terms of \( p_i \) and \( L \):

\[
S_{it}(q) = \ln p_i^{-1} + \ln L^{-1}.
\]

For basic information theoretic and Bayesian notions and notations see, e.g., the MacKay (2003) textbook. The framework given here was used in a CMB context by Farhang et al. (2011). For most the \( q_k \) are the amplitudes of the ordered eigenmodes for XeMs or eXeMs. Generally, the fluctuations in the standard cosmic parameters from their maximum likelihood values are included along with these eigen parameters. We shall assume the prior distribution of the parameters to be uniform in the \( q_k \).

The expansion of \( S_{it} \) to quadratic order is the basic perturbative approach used throughout this paper, leading to a Gaussian

\[
S_{it}(q) \approx S_{t,m} + q^T F q / 2
\]

terms of the Fisher matrix and the information action minimum \( S_{t,m} = -\ln(p_i L_{\text{max}}) \).

The posterior Shannon entropy is related to the final-state ensemble average of the information action and the evidence:

\[
S_t \equiv \langle \ln p_i^{-1} \rangle_t = \langle S_{it}(q) \rangle_t + \ln E.
\]

For the quadratic order expansion it is

\[
S_t \approx \frac{1}{2} \text{Tr} \ln F^{-1} + \frac{N}{2} \ln(2\pi) + \frac{1}{2} \text{Tr}(\langle qq^\dagger \rangle F)
\]

\[
= \frac{1}{2} \text{Tr} \ln F^{-1} + \frac{N}{2} (\ln(2\pi) + 1).
\]

The second line follows from the first since the correlation matrix of the \( q \) is \( \langle qq^\dagger \rangle = F^{-1} \). The associated evidence involves the information action minimum, \( \ln E \approx S_t - S_{t,m} - (N/2) \).

The entropy associated with mode \( n \) is

\[
S_n \equiv -\frac{1}{2} \ln f_n + (1 + \ln 2\pi) / 2 = S(\leq n) - S(\leq n - 1).
\]

It is a finite difference of the total entropy of the first \( n \) modes in the eigen hierarchy,

\[
S(\leq n) = \frac{n}{2} (1 + \ln 2\pi) - \frac{1}{2} \sum_{k=1}^{n} \ln f_k
\]

and

\[
\langle s \rangle_n \equiv S(\leq n) / n
\]

is the associated mean entropy per mode. Figure 11 shows how the relative entropy \( S_n - S_1 \) and the mean entropy \( \langle s \rangle_n - S_1 \) = \( S(\leq n) / n - S_1 \) grow with \( n \) for the modes derived from the localized Gaussian expansion. We also plot two versions of “white-noise” entropy:

\[
S(wn, \leq n)(\sigma^2) \equiv n(\ln \sigma + (1 + \ln 2\pi) / 2),
\]

\[
\text{mean – variance } \sigma^2 = \sum_{k=1}^{n} f_k^{-2} / n,
\]

\[
\text{mean – weight } \sigma^2 = \left[ \sum_{k=1}^{n} f_k^2 / n \right]^{-1}.
\]

These are entropies maximized subject to the constraint that we only have knowledge of the integrated \( \sigma^2 \), whereas \( S(\leq n) \) is the maximized entropy given knowledge of the full spectrum \( \{ f_k^{-1} \} \). The mean-variance white noise lies above \( S(\leq n) \) and the mean-weight white noise lies below. The mean-weight behavior is dominated by an \( \ln(n) \) rise, since the total weight of modes below \( n \), \( \ln \sum_{k=1}^{n} f_k^2 \), quickly approaches a constant, reflecting the dominance of the high-weight eigenmodes in the sum.

We first discuss why we do not use the traditional evidence ratio often used in Bayesian theory to decide if a new parameter
$q_n$ should be added. The log-evidence difference for the addition of $q_n$ is

$$
\Delta \ln \mathcal{E}_n = \ln \mathcal{E}(\leq n) - \ln \mathcal{E}(\leq n - 1) = S_n - 1/2 - \Delta S_{\ln}.
$$

This requires evaluation of the change in the information minimum. It also has the usual disadvantage of depending upon the $f_k$ measure. Although using eigen parameters ensures the same dimensionality for the different $f_k$, it does not fully remove this re-parameterization ambiguity since there can be a $k$-dependent scaling. (In fact, we have usually renormalized our $f_k$ so that the associated eigenmodes $E_k(z)$ have unit norm upon $z$-integration.)

Our preferred approach for hierarchy truncation is to use suitably defined entropy differences. In particular, we wish to set a threshold control on the injection entropy, $\delta S_{ij,n} \equiv S_n - \langle s \rangle_n$, the entropy from adding mode $n$ relative to the mean entropy from all $\leq n$ modes. It is related to the relative increase in phase-space volume $V(\leq n) = \exp(N(\leq n)/n/2) = \exp(n\langle s \rangle_n - 1/2)$ associated with mode additions:

$$
\ln[V(\leq n + 1)/V(\leq n)] = S_{n+1} - \langle s \rangle_n.
$$

We choose $S_{ij,n}$ instead because it is zeroed out for mode one, but $S_{ij,n}$ quickly approaches $S_n - \langle s \rangle_n$. For example, if we impose a $\Delta S \sim 1/2$ threshold in Figure 11 on the CVL XeM case, we would use only one mode, whereas $\Delta S \sim 1$ picks up about 5, $\sim 3/2$ harvests about 10, and 2 gives about a dozen. Similar tales can be told for the eXeM CVL case and for both Planck+ACTPol forecasts. Another more erratic measure is relative injection jumps, which is nearly $S_{n+1} - S_n$. In Figure 11, the negative is plotted for clarity of presentation. Either reading off from the figure, or using the lists of errors in Tables 1 and 2, the sample threshold $\Delta S \sim 1/2$ again yields only a mode or two.

The fluctuating nature of $S_{n+1} - S_n$ implies that we can use it to split the modes into groups of similar information content which arise by thresholding it. Thus, for a chosen threshold value all the modes between two successive boundaries, where $\Delta S$ exceeds a certain value, are considered as one mode group. If a mode is selected to be included in the analysis by, say, sharp-thresholding the injection entropy, it is logical that all of its co-modes be included, which is akin to softening the threshold. The groupings found with $\Delta S \sim 1/2$ imposed upon $S_{n+1} - S_n$ create boundaries at one mode, five modes, and so on. These are, not surprisingly, similar to mode numbers obtained as we move the threshold on injection up; hence that criterion can be used instead to define mode groups.

Although these entropy difference criteria imply that relatively little additional information is gained by including more than a handful of higher modes, in real data analysis the situation is subtler, with other criteria important to consider. For example, depending on how close the assumed model is to the true underlying history, our measurements of standard cosmic parameters might be biased. In that case one would like to add enough modes to remove the bias, sequentially checking if the recovered values of the standard parameters are robust against introduction of the next eigenmode. A reasonable strategy is to add one mode group at a time to the analysis until the biases are removed. In the next section, we show how varying the mode number cutoff affects our results, roughly following this grouping procedure.

### 3.7. Perturbation Reconstruction: Eigenmodes as a Complete Basis

Any function $X_e(z)$ (in the redshift range under consideration) can be expanded in terms of these XeMs unless it has highly localized features compared to the highest frequency present in the basis functions or to the width of the bumps in the case of localized modes. That is because the XeMs are just linear combinations of the original basis functions, and thus cannot have frequencies higher than the maximum frequency present in the basis functions. On the other hand, as is clear from Figure 5, strongly localized features in possible perturbations to recombination history are not constrained with CMB data sets. Therefore, the lost features of an ionization model via expansion by these eigenmodes are not measurable even if modes with higher frequencies are included in the analysis. In other words, the XeMs serve as a complete basis for the expansion of constrainable features in the possible perturbations in the recombination history.

To demonstrate the reconstruction of perturbations using the XeMs we choose two physically motivated ionization perturbations, one associated with physical corrections to the recombination process (Chluba & Thomas 2011, hereafter CT2011) and the other due to energy injection coming from a model of DM annihilation (using the description of Chluba 2010).

#### 3.7.1. Standard Recombination Corrections

The modification to $X_e$ corresponding to CT2011 is shown in the top left panel of Figure 12 (black solid line). This correction should be added to the $X_e$ from the original version of RECFAST (or the $X_e$ from RECFAST v1.4.2 setting $\text{HeSwitch} = 0$). At high redshifts one can see the effect of accelerated helium recombination caused by absorption of photons in the Lyman continuum of hydrogen. During hydrogen recombination the corrections are caused by detailed radiative transfer effects as well as two-photon and Raman scattering events. The freezeout tail is slightly higher than obtained with RECFAST because of deviations from statistical equilibrium in the angular momentum sub-states. We note that with RECFAST v1.5 a large part of all these corrections can be accounted for; however, these corrections are not explicitly modeled using a physical description but have been judged to reproduce the results obtained with detailed recombinaton codes.

We project this $\delta \ln(X_e)$ on the 160 XeMs constructed from perturbations in the recombination history in the form of Gaussian bumps and with the perturbation parameter being $\delta u(z) = \delta \ln(X_e)$ for the CVL case with $\ell_{\text{max}} = 3500$, as described in Section 3.1. Figure 12 compares the reconstructed perturbation for three cases with a different number of XeMs included. First, note that by including only 160 XeMs the original perturbation is practically fully recovered. If only the 15 most constrained modes are included, the helium correction ($z \sim 1800$) and also hydrogen correction around $z \sim 1100$ are well restored while for lower $z$ regions higher modes are required. The reconstruction by six XeMs, however, is most sensitive to variations around $z \sim 1100$ and cannot tell much about the helium correction. The projection coefficients for the first six XeMs are shown in Table 4. For this particular model of corrections to the perturbation scenario we see that the XeMs 1, 3, and 6 are strongly dominant among the first six modes.
Figure 12. Reconstruction of two physically motivated Xe perturbation scenarios on (different number of) XeMs generated with Gaussian bumps (top) and the relative difference in the temperature power spectrum between the reconstructed perturbations and the full corrections (bottom). Right: the perturbations come from deviation from physical corrections to the recombination process (CT2011). Here the perturbation parameter is $\delta \ln(X_e)$. Left: the perturbations are due to a model of dark matter annihilation. As the perturbation parameter we used $\delta \ln(X_e + 0.01)$ to better accommodate for the freezeout perturbation. A case with $\delta u(z) = \delta \ln(X_e)$, i.e., $\sigma = 0$, is shown for comparison.

(A color version of this figure is available in the online journal.)

Table 4

| XeM               | 1  | 2  | 3  | 4  | 5  | 6  |
|-------------------|----|----|----|----|----|----|
| CT2011 ($\sigma = 0$) | $-0.32$ | $0.08$ | $0.16$ | $0.02$ | $-0.09$ | $0.25$ |
| DM annihilation ($\sigma = 0.01$) | $-0.31$ | $-0.30$ | $0.46$ | $-0.14$ | $0.33$ | $0.88$ |

Note. Note the different values of $\sigma$ used for the two cases.

The lower left panel in Figure 12 illustrates the relative difference between the temperature power spectrum for the reconstructed perturbations and the original full corrections. We see that with only six modes the error in the recovered $C_\ell$ is less than 0.1% although the difference with respect to the SRS is $\sim$4% at high $\ell$. Remembering that the changes in the $C_\ell$ due to the full corrections are about a few percent, this shows that the main corrections to the CMB power spectra can be captured by just introducing a small number of modes. The CMB data indeed are not very sensitive to all the details in the freezeout tail of recombination and during helium recombination, unless prior knowledge renders uncertainties at $z \sim 1100$ very small. As we see below, part of the corrections from higher modes are compensated for by biasing the XeMs included in the analysis.

3.7.2. Dark Matter Annihilation Scenario

As the second example we chose the perturbations arising from a model of DM annihilation. It was computed using the description of Chluba (2010) with an annihilation efficiency $f_{\nu{DM}} \sim 2 \times 10^{-24} \text{eV s}^{-1}$. The difference with respect to RECFAST is shown in the right panel of Figure 12. In contrast to the previous case, the perturbations here are not concentrated around the maximum of differential visibility but are most significant at lower redshifts. Therefore, for the decomposition of the DM perturbations we choose $\delta u(z) = \delta \ln(X_e + 0.01)$.
To allow a better recovery of the relatively large perturbations in the freezeout tail without the need to include too many modes. This procedure can be interpreted as placing a strong prior on (physically) expected changes in the freezeout tail.

The top right panel of Figure 12 shows the reconstructed perturbation including three different numbers of XeMs. Here the recovered curve becomes very close to the original perturbation by including the first 15 XeMs, while six XeMs have a poor recovery of the low-z part. Note that the plots are illustrating \( \Delta \ln(X(z)) \) although the XeMs and thus the decomposition of the perturbation are all performed with \( \Delta \ln(X_c + 0.01) \).

For comparison the reconstruction of the perturbation with \( \Delta \mu(z) = \Delta \ln(X_c) \), i.e., with \( \sigma = 0 \), and with 15 XeMs taken into account is also shown. As expected, this reconstruction is much poorer compared to the previous case with \( \Delta \mu(z) = \delta \ln(X_c + 0.01) \) due to its lack of coverage of corrections in the freezeout tail. This demonstrates that when there is prior knowledge in favor of the freezeout tail of recombination being affected, a parameterization with \( \sigma > 0 \) should be used in the analysis. However, it is still correct that the main signal is produced by the modifications close to \( z \sim 1100 \), even if the freezeout tail apparently has the largest deviation from the SRS. This is why the first few mode functions for \( \sigma = 0 \) do not have any strong low-redshift tails. The eigenvectors naturally order the perturbations in the strength of the associated change in the CMB power spectra, as explained in Section 3.1. This point is visible from the lower right panel where the \( C_\ell \) difference is plotted for reconstructed perturbations with different number of modes included compared to the full perturbations. Similar to the previous case, these differences are several times smaller than the changes in the \( C_\ell \) caused by this model of DM annihilation, again meaning that these few modes can well capture the constrainable features of the perturbations.

Also, if we look at the decomposition of the recombination correction into the first six XeMs (see Table 4) we see that they all have comparable contributions. This seems reasonable if we remember that the mode functions, despite being weighted toward the low-redshift part, still have a significant component at high redshift which need to be canceled out to recover this pattern of perturbation with its low-redshift modification. Therefore, the neighboring modes have the same order of amplitude to properly cancel out the high-redshift perturbations. This implies that the distribution of the mode amplitudes can in principle hint to the type of perturbation involved. However, a detailed analysis of this kind requires a model selection study. For example, in the data analysis, one could treat \( \sigma \) as a hyperparameter and estimate its best-fit value at the same time as the corresponding perturbation eigenmodes and the standard cosmic parameters. This allows us to choose the most preferred model among the class of all models parameterized by \( \sigma \).

### 4. MEASURING THE AMPLITUDES OF PERTURBATION EIGENMODES FOR SIMULATED DATA

Having constructed the eigenmodes, their amplitudes can now be considered as additional parameters to be plugged into CosmoMC, “a Fortran 90 MCMC engine for exploring cosmological parameter space”. In this section, we investigate how well the most constrained XeMs and eXeMs can be measured by simulated data. To study the impact of these new variables on the standard parameter estimation, we first consider the case in which the data are both simulated and analyzed using the SRS (Section 2.1). We then study the case for which the effects of physical corrections to the recombination history (CT2011, see Section 3.7) are included in the constructions of the mock data, but are neglected in the fiducial recombination model used in the analysis. Here the question is how well the eigenmodes compensate for the deviations from the fiducial model and how much the data are telling us about the amplitudes of the modes. We also briefly discuss how the eigenmodes should be used in a more general case where little prior knowledge about the recombination perturbations is available.

#### 4.1. Case 1: The Standard Recombination Scenario

As the first example we choose the fiducial recombination model (here the SRS) to be identical to the ionization history used in the simulation of the data. We ran CosmoMC to estimate the best-fit values and errors of the six standard parameters together with those for the perturbation eigenmodes. We tried the two sets of eigenmodes described before: the first five XeMs and the first six eXeMs. The number of modes in each case was chosen in rough agreement with the mode cutoffs described in Section 3.6 (more precisely with \( \Delta S_i = S_{n+1} - S_n = 1/2 \)). The simulations were carried out for a CVL experiment.

One expects no detection of eigenmodes since the fiducial model for \( X_e \) and the underlying model used to simulate data are the same, as verified by Figures 13 and 14. Also, by construction there is almost no visible correlation between the measured parameters for the eXeMs, at least sufficiently close to the best-fit model where the assumptions of the Gaussianity for the likelihood surface approximately holds. However, Figure 13 indicates that the XeMs become partially correlated with each other, although by construction these were initially uncorrelated. The reason is that the standard parameters were held fixed during the process of XeM construction, but now that they are allowed to vary, their degeneracy with the XeMs induces correlations. These new correlations lead to larger errors than those deduced from the simple Fisher analysis (Table 1 compared with Table 5) and can also change the rank ordering of the modes, e.g., the error on XeM 2 is smaller than XeM 1 (Table 5).

The standard parameters remain unbiased, as the model used for simulating data and the theoretical model used in the analysis were the same. This is no longer true once recombination corrections to the SRS are added (see Figure 17). However, the correlations of the eigenmodes with some of the standard parameters increase the errors of the standard parameters. From Figure 15 we see that among the standard parameters, \( \Omega_m h^2 \), \( n_s \), and \( A_s \) are the ones most affected by the introduction of the eigenmodes into the analysis. This can be understood by noting the relatively high degeneracy between these parameters and some of the eigenmodes. The most evident one is the

\[ \text{Table 5} \]

| XeM | 1  | 2  | 3  | 4  | 5  |
|-----|----|----|----|----|----|
| 1σ  | 0.046 | 0.030 | 0.057 | 0.088 | 0.086 |

13 http://cosmologist.info/cosmomec/

14 We confirmed this statement by running MCMC with non-varying standard parameters.

15 It should also be noted that the correlations between the standard parameters themselves also change when the eigenmodes are introduced.
correlation of \( n_s \) with the first XeM which by changing the width of the visibility function leads to a tilt in the power spectra (compare Figures 16 and 10). For the case of \( \Omega_b h^2 \) and \( A_s \), it is harder to give a visual interpretation. \( \Omega_b h^2 \), leading to both tilt changes and oscillations in the \( C_T \), correlates with most of the first five XeMs (the highest being with XeM 1), while \( A_s \), being an amplitude multiplier, mainly correlates with XeM 1. These correlations between the standard parameters and the eigenmodes emphasize the fact that uncertainties in the recombination scenario in particular undermine our ability to measure the precise values of \( n_s \) and \( \Omega_b h^2 \) (see, e.g., Shaw & Chluba 2011). Also note that the changes in the error bars of the standard parameters are actually practically independent of which set of eigenmodes are used (Figure 15). This suggests that in terms of the standard parameter estimation, the use of XeMs or eXeMs should not lead to vastly different results in the parameter estimation. However, the perturbations are measured to higher accuracy with the eXeMs (Table 2) than XeMs (Table 5) especially if only a few modes are included in the analysis. Therefore, as long as only CMB data are used, the eXeMs are the more appropriate choice of eigenmodes.

Finally, we studied how much the presence of perturbations to recombination could affect our ability to determine the precise value of \( Y_p \). The abundance of helium affects the CMB anisotropies mainly because more helium implies fewer free electrons during hydrogen recombination. Consequently, \( Y_p \) should also couple significantly to the perturbation eigenmodes. We therefore performed simulations in which \( Y_p \) was also allowed to vary. The analysis was performed with three and five XeMs in the Planck-ACTPol-like and CVL cases, and with six eXeMs for the simulated CVL data. Table 6 compares the 1σ error bars on \( Y_p \) in these cases. We see that for the CVL case similar number of XeMs and eXeMs used as the eigenmodes lead to similar constraints on helium abundance. However, a Planck-ACTPol-like observation gives a few times larger error due to lack of very high sensitivity to very small scales, although fewer XeMs compared to the CVL case have been used.

### 4.2. Case 2: A Perturbed Recombination Scenario

As the second example of parameter estimation and perturbation reconstruction, we simulate data assuming the recombination calculation of CT2011 (Figure 12), while we take the fiducial model to be as of RECFAST v1.4.1 or older (equivalent to SRS with He_{switch} = 0 to remove the helium correction which has been assumed as part of the perturbations in the data). The purpose here is to find out how well the biases in the standard parameters due to this lack of knowledge about the physical corrections can be removed by including the perturbation eigenmodes, and whether or not, data can reconstruct part of the true recombination history.

Constructed from CosmoMC chains for a CVL experiment, Figure 17 illustrates the two-dimensional contours of some of the standard parameters. The large biases in the estimated values of the parameters when only the six standard parameters are measured are due to the mismatch between the ionization history in the theoretical model and the data. Here, only contours for parameters with the largest biases are shown. See also Shaw & Chluba (2011). To compensate for this mismatch we separately add to the parameters the two different sets of the eigenmodes, the XeMs and eXeMs, as the new parameters.

As Figure 17 demonstrates, without the eigenmodes the bias in, e.g., \( \Omega_b h^2 \) and \( n_s \) is about 5σ and 7σ, while adding the eigenmodes eliminates the bias at the cost of increased error bars.
shown by the black diamond. To the case with no eigenmodes included. The input value of the parameters is by an ideal experiment in the presence of five (six) XeMs (eXeMs) compared (the difference in measured uncertainty of \( \tau \) with and without the eigenmodes is rather small). Our computations also indicate that with the XeMs as the eigenmodes and for a CVL observation, the minimum number of modes required to remove the bias from these standard parameters is six. However, we included eight modes in the analysis to take into account the mode-selection criterion of Section 3.6 (determined by the relative injection jumps).

We also observe that the recovered values for the amplitude of the XeMs are biased (compared to the theoretically expected values from direct projection on the XeMs, Table 4) and change by varying the number of modes included in the analysis. That is due to the correlation of the XeMs in the presence of the standard parameters, and the fact that not all XeMs are included into the parameter estimation. As a result, parts of the perturbation that project on the neglected higher XeMs leak into the lower XeMs. The bias in the measured XeM amplitudes is similar to the bias in the standard parameters when there are no eigenmodes in the analysis, but with a much lower significance.

For the same reason, the errors on the XeMs also change when the number of modes included in the analysis changes. However, as mentioned before, due to the low significance of the perturbation detection for most of the XeMs this is not as important as for the main cosmic parameters. For the CVL simulations with six and eight XeMs included, we see that the most significant contribution comes from the first mode (respectively \( \mu_1 = -0.23 \pm 0.05 \), \( \mu_1 = -0.18 \pm 0.04 \)) while the other modes are consistent with zero. This is also true for a Planck-ACTPol-like case, which we come to shortly, where the first XeM is measured to be \( \mu_1 = -0.22 \pm 0.06 \) and \( \mu_1 = -0.24 \pm 0.12 \) for one and three XeM measurements.

If instead eXeMs are used as the perturbation eigenmodes, our computations show that at least 10 modes should be added to get rid of the bias for a CVL case. However, as a test case, we tried including the best 20 eXeMs (see Figure 17). We also found that although the errors on the standard parameters keep increasing by adding more eXeMs to the analysis up to around the 10th mode (which is required to remove the bias), it stays more or less the same afterward. This suggests that in terms of the constraints on the standard parameters, we do not lose much by increasing the number of eXeMs. Besides, including more eXeMs does not affect the measurement of the previously included eigenmodes, as they are by construction uncorrelated (in the presence of standard parameters). Including more eXeMs, on the other hand, makes the reconstructed perturbations closer to the input model of perturbations (as in Figure 12). However, as the errors of modes increase by going to higher orders, the error on the reconstructed curve increases. We address this point in the next section.

Among the first 20 eXeMs for a CVL experiment, the modes with the most significant contributions (i.e., with at least 1\( \sigma \) detection) are \( \mu_2 = 0.11 \pm 0.02 \), \( \mu_3 = 0.10 \pm 0.03 \), \( \mu_4 = -0.31 \pm 0.16 \), and \( \mu_{11} = -0.36 \pm 0.24 \) (compare to their theoretical prediction from direct projection of the perturbations on the eXeMs: \( \mu_2 = 0.14 \), \( \mu_3 = 0.10 \), \( \mu_4 = -0.33 \), and \( \mu_{11} = -0.39 \)). The reason that the recovered value, though close, is not exactly the same as the forecast is that the assumption of the Gaussianity of the distributions of the eXeMs and the standard parameters is only approximate. Also, the eigenmodes have been slightly smoothed in the construction process, which may cause numerical inaccuracy and induce slight correlation between the smoothed modes. By comparing the theoretical values of projection of the perturbation on the eXeMs and their forecast errors (from Fisher analysis) we do not expect any perturbation detection after eXeM 11.

Figure 18 shows similar contours but for a simulated Planck-ACTPol-like observation. For the analysis we used the eigenmodes (both eXeMs and XeMs) constructed with the Planck-ACTPol simulated noise. The results from the two sets of eigenmodes are very similar. For both XeMs and eXeMs, one mode was sufficient to remove the bias (\( \mu_1 = -0.22 \pm 0.06 \) and \( \mu_1 = -0.20 \pm 0.06 \), respectively). This happens to be in agreement with the cutoff mode for the XeMs while with the eXeMs...
Figure 16. Derivatives of the $C^{T,E}_\ell$'s with respect to some of the standard parameters.
(A color version of this figure is available in the online journal.)

Figure 17. Contours of some of standard parameters for CT2011 case, with eight XeMs in one case and 20 eXeMs in the other case included in the analysis, compared to a case where no perturbation eigenmodes (of any kind) have been included (the solid red curves). The simulations are performed for a CVL experiment. The input value of the parameters is shown by the black diamond.
(A color version of this figure is available in the online journal.)

Table 6

|                   | CVL (std)     | Planck-ACTPol (std) | CVL (std + 5 XeMs) | CVL (std + 6 eXeMs) | Planck-ACTPol (std + 3 XeMs) |
|-------------------|---------------|---------------------|-------------------|---------------------|-----------------------------|
| $Y_p$             | 0.240 ± 0.0016| 0.240 ± 0.006       | 0.239 ± 0.005     | 0.240 ± 0.004       | 0.238 ± 0.017               |

the second mode should also be included. The lower number of modes required for the Planck-ACTPol-like case compared to the ideal experiment is expected due to higher sensitivity of the data in the latter to deviations from the underlying $X_e$ history. We also tried three modes, with no significant detection of the new modes, while the error on the XeM 1 increases by a factor of two.

4.3. Trajectories

In this section, we investigate the reconstruction of the $X_e$ perturbations using the simulated data to illustrate the corresponding uncertainty at different redshifts. The left plot in Figure 19 shows the redshift interval covered by 500 $\delta \ln X_e$ trajectories corresponding to an ideal observational case with eight XeMs included, for the CT2011 model. The color indicates the number of trajectories passing through each $(z, \delta X_e/X_e)$ bin, normalized to one at each redshift snapshot. The trajectories clearly show deviations from the SRS, slowly morphing into the correction obtained by CT2011 (the cyan curve). However, the recovery is not perfect, as the model of CT2011 has non-zero (and relatively significant) projection on higher XeMs which are not well constrained by data and therefore were not included into the analysis. Most obviously, corrections to helium recombination are not captured well when using only the first few XeMs. These trajectories do not recover the analytical projection of the CT2011 corrections on the first eight XeMs very well either. The reason, as discussed before, is that the correlation of the XeMs induced by the standard parameters draws some contribution from the higher absent modes, which biases the measurement of the first few XeMs included in the measurement.

To test this impact of higher, excluded modes on the recovered (low XeM) trajectories, we ran simulations with the data that only accounted for the contributions from these low modes. As expected, in the absence of higher modes in the data, the measured XeMs were non-biased and thus the highest probability region of the trajectories covered the $\delta \ln X_e$ curve of the input model.

Although our basic target is $X_e$ recovery, the relevant space for determining how well we have done is that of the CMB data, reduced to the power spectra, $C^{TT}_\ell$ and $C^{EE}_\ell$. The central and right panels of Figure 19 show the $\delta C^{TT}_\ell/C^{TT}_\ell$ and $\delta C^{EE}_\ell/C^{EE}_\ell$ trajectories, where $\delta C^{X}_\ell/C^{X}_\ell = (C^{X}_\ell - C^{X,\text{fid}}_\ell)/C^{X,\text{fid}}_\ell$ and $C^{X,\text{fid}}_\ell$ is the fiducial power spectrum without any perturbations. The transformation from $X_e$ trajectories to $C_\ell$ trajectories shows a much tighter band around the input signal. This is a visual
confirmation of the point that some features in the $\delta \ln X_e$ which make the $X_e$ trajectories thick do not leave a measurable imprint on the $C_\ell$. Note that there are small residual oscillations, i.e., in the difference between the recovered trajectory and the input power spectrum. They coincide with the peaks and troughs of the $C_\ell$ curves for both $TT$ and $EE$ (which is out of phase with $TT$). One source for the oscillations seems to be the eigenmode truncation, as we will see later. Using only a limited number of the modes in the analysis causes the non-$X_e$ cosmic parameters to try to match the injected $X_e$ perturbations. There is also an issue of accuracy of the $C_\ell$ code for some of the distortions.

Figure 20 similarly shows the two-dimensional histograms of trajectories for the case with the first 10 eXeMs included. Around the maximum of the Thomson visibility function the $X_e$ reconstruction is slightly stronger and less fuzzy than in Figure 19, with part of the helium recombination correction being recovered. The improvement in the reconstruction is because for the computation of these eigenmodes their correlation with the standard parameters have been optimally taken care of. In contrast, the XeMs used in the previous case are non-optimal if no strong additional priors can be placed on the standard parameters, leading to confusion in the errors and the rank ordering of the modes. Figure 21, constructed with 20 eXeMs included, shows that the oscillation effect mentioned above around the input $C_\ell$ signal is diminished (and also partially swamped by the slightly higher dispersion around the input curve) for the 20 mode compared to the 10 mode case. We also see that including a higher number of modes does not necessarily lead to better $X_e$ recovery.

Similar trajectories for a Planck-ACTPol-like experiment are shown in Figures 22 and 23, with three XeMs and three eXeMs as the eigenmodes, respectively. The trajectories for the XeMs are more widely spread and blurred due to experimental noise. The eXeMs perform slightly better. However, the overall reconstruction is clearly lacking detailed agreement with the full recombination correction of CT2011. In particular, most of the modification during helium recombination is not captured, as the corresponding signals can only be picked up with higher modes, which in the considered case are not constrainable at a significant level. In the $\delta C_{\ell}/C_{\ell}$ plots of these Planck-ACTPol-like cases, there is a small disagreement at high multipoles between the theoretical curve and the highest probability region of the chains. That is mainly due to mode truncation at a relatively low mode number, i.e., three. We tested this by including eight modes and, as expected, observed a wider spread around the input signal with the disagreement diminished.

Although we do not plot the equivalent $\delta C_{\ell}/C_{\ell}$ for the DM case discussed in Section 3.7.2, very similar plots result, namely, good recovery of the power spectra with a dispersion around the input perturbation signal.

### 4.4. Beyond Small Perturbations

In this paper, it was explicitly assumed that the model best explaining the ionization fraction (or the true model underlying the ionization history) is only slightly different from our fiducial model, justifying our choice of parameter $\delta \ln X_e$. Therefore, the eigenmodes constructed for the fiducial model are also very close to the eigenmodes for the perturbations to the true $X_e$, the corrections to the eigenmodes arising from the
Figure 20. Similar to Figure 19 but with the first 10 eXeMs.
(A color version of this figure is available in the online journal.)

Figure 21. Similar to Figure 19 but with the first 20 eXeMs. As this figure demonstrates, including a higher number of modes does not necessarily lead to better $X_e$ recovery. Here the recovered $X_e$ becomes noisier compared to the case with only 10 modes included, while the $C_\ell$ trajectories do not change significantly except for the diminished oscillations around the input model, as discussed in the text.
(A color version of this figure is available in the online journal.)

Figure 22. Similar to Figure 19 but for a Planck-ACTPol-like experiment and with only three XeMs taken into account.
(A color version of this figure is available in the online journal.)

Figure 23. Similar to Figure 20 but for a Planck-ACTPol-like experiment and with only three eXeMs taken into account.
(A color version of this figure is available in the online journal.)
difference between the fiducial and true $X_e$ model being only of second order. Under this assumption, a one-step search for the best-fit parameters suffices to extract the available relevant information from the data, provided that the minimum required number of modes are included in the analysis. Finding that the minimum number of required modes can by itself involve several parameter estimation steps in parameter spaces with different dimensions, the criterion being that the best-fit values for the standard parameters stop changing. That is what was done in the examples in this work (Section 4.2) to illustrate how the method works.

However, if the fiducial model is very far from the true $X_e$ history, such that the eigenmodes are expected to be affected at a significant level, an iterative approach toward finding the best modes with their associated amplitudes and errors is required: starting with our best guess for the fiducial model, we estimate its deviation from the true ionization history using the data set available and the eigenmodes constructed based on this fiducial model. We then update the model by adding to it the measured deviations in the eigenmodes (and the standard parameters, if required). This process is repeated until the convergence of the model and its eigenmodes.

However, current constraints seem to indicate that such an iterative procedure is not necessary within the standard picture. For example, as shown by Shaw & Chluba (2011), the recombination corrections of CT2011 are readily incorporated using one calibrated redshift-dependent correction function relative to the original recombination model of Seager et al. (1999). Even for CVL errors a second update of the correction functions leads to minor effects. Nevertheless, if something more surprising occurred during recombination, an iterative approach might be required.

5. CONCLUSION AND DISCUSSION

CMB data today are becoming so precise that small modifications in standard ionization history are important. This impressive progress not only implies that measurements of the main cosmological parameters are becoming increasingly accurate but also means that remaining uncertainties in the recombination dynamics, e.g., caused by neglected standard or non-standard physical processes, should be quantified. In this work we discuss a novel approach to constrain this remaining ambiguity with future CMB data. We performed a principal component analysis to find parameter eigenmodes that can be used to describe uncertainties in the ionization fraction. We constructed $X_e$ eigenmodes over the redshift range of [200, 3000], performing several consistency checks to prove the correctness of our method. This approach automatically delivers a hierarchy of mode functions that can be selected according to their error and then are added to the standard cosmological parameters when analyzing CMB data.

Due to the strong CMB signal imprinted by hydrogen recombination, the most constrained modes are mainly localized around $z \sim 1100$, with some extensions to lower and higher redshift regions (see Figures 5 and 7). This emphasizes that CMB data are very sensitive to small changes during hydrogen recombination, while details of helium recombination or small changes in the freezeout tail are hard to constrain, unless strong priors on the reliability of the hydrogen recombination model are imposed. With the method described here it is possible to construct mode functions for different experimental situations, also folding in prior knowledge on the recombination history using appropriate weight functions and fiducial $X_e$ models. For example, if there are physically motivated and experimentally supported hints toward (significant) changes in the freezeout tail of recombination, e.g., due to energy injection from DM annihilation, we propose a parameterization which weights the low-redshift part more strongly (see Figure 4).

After we completed this work, we received a preprint (Finkbeiner et al. 2012) that investigated the use of CMB data to constrain details of energy injection scenarios related to decaying or annihilating particles. They also used parameter eigenmodes, but these were constructed based on an energy release history which is in our language akin to the imposition of a strong prior on the recombination dynamics around $z \sim 1100$ and a focus on the freezeout tail of the recombination.

We applied the method to different simulated data sets with the aim to assess how well future CMB experiments will be able to constrain modifications to the standard recombination scenario. (Current WMAP plus ACT and SPT data will provide only relatively weak constraints, but Planck plus ACTPol and SPTPol will considerably improve the situation.) As a working example we used the refined recombination calculations of CosmoRec. For simulated CMB data sets corresponding to Planck-ACTPol-like experiments we found that the first three eigenmodes can be rather well constrained. The addition of these modes allows us to compensate for the measurable differences between the fiducial old $X_e$ model as given by RECFAST to the new recombination history computed with CosmoRec, and thus partially reconstruct the true $X_e$ history, without actually directly using the recombination corrections in the analysis. However, because the first few mode functions are strongly localized around $z \sim 1100$, details during helium recombination and in the freezeout tail are not captured (Figure 23). The addition of the first three eigenmodes is sufficient to remove the biases in the standard parameters, however, at the cost of increased error bars. We also show that for CVL-limited experiments up to $t \sim 3500$ up to 10 modes might be constrainable, in this case allowing us to pick up part of the details during helium recombination (Figure 20).

The significance of the detection of any perturbation obviously depends on the underlying ionization history of the real data. In the specific CosmoRec example for Planck-ACTPol-like experiments, all three modes but the first one are consistent with zero. A significant source for large errors on the eigenmodes is their correlation with the standard parameters. If tight constraints are imposed on the standard parameters by non-CMB experiments such as Baryon acoustic oscillations or supernova data, the errors on the eigenmodes will be correspondingly reduced. Comparing the first rows of Table 1 (where all standard parameters are held fixed) and Table 2 (where all standard parameters are being marginalized over) illustrates the effect of this correlation in the extremes.

This also shows how important one’s knowledge of how well elements of recombination that are known, expressed through prior probabilities, will be. If the uncertainty in the ionization history during hydrogen recombination can be reliably reduced by other methods then the sensitivity to small perturbations at higher or lower redshifts is enhanced. We note that measurements of the cosmological recombination radiation (e.g., see Chluba & Sunyaev 2006a; Sunyaev & Chluba 2009) could in principle provide an alternative way of constraining the recombination dynamics in the future. In particular, the recombination radiation could exhibit significant features if something more unexpected occurred during different cosmological epochs (e.g., see Chluba & Sunyaev 2009; Chluba 2010).
In principle, for a complete study of ionization history, late reionization should also be included in the analysis. Ambiguities in the low-redshift part of the ionization history may affect the measurements of high-redshift perturbations and vice versa. However, the main signal from the reionization epoch is measurable from the very-large-scale CMB polarization, and the high-redshift perturbations of $X_e$ affect anisotropies with smaller angular scales. Therefore, the signals from these two regions are rather uncorrelated. A more complete analysis for the whole ionization history or where different parts of it are considered simultaneously is for future work.

An aspect requiring a decision when analyzing real data is the choice of parameterization. For most of this work we weighted the perturbations in $X_e$ by the fiducial history. If, for example, the recovered perturbations point toward modifications in the freezeout tail of recombination, or if there is a strong belief that no sign of significant deviations around the maximum of visibility are present, an alternative parameterization that allows better reconstruction of the tail can be constructed, using appropriate weight functions that quantify our belief in the underlying fiducial model.

As discussed in Section 4.4, our semi-blind XeMs are designed to only probe small perturbations about the fiducial model $X_e^\text{fid}$. When it comes to real CMB data analysis, iterations of $X_e^\text{fid}$ may be required to ensure no leftover bias remains. We look forward to the application of iteratively improved eigenmodes to the coming high-resolution CMB data from Planck, ACTPol, and SPTPol.

We thank Doug Finkbeiner, Tongyan Lin, and Olivier Doré for useful discussions and Eric Switzer for his comments on the text. Support from NSERC and the Canadian Institute for Advanced Research is gratefully acknowledged.

APPENDIX A
LOCALIZED BASIS FUNCTIONS

To expand perturbations to the ionization scenario, three sets of localized mode functions have been considered in this paper: Gaussian and triangular bumps and $M_4$ splines. For the Gaussian and triangular bumps, we define the $i$th basis function centered at redshift $z_i$ and having width $\sigma_i$ by

$$\varphi_i(z) \propto \exp \left( -\frac{(z - z_i)^2}{2 \sigma_i^2} \right)$$

(A1)

for the Gaussian case and by

$$\varphi_i(z) \propto \begin{cases} 1 - \frac{|z - z_i|}{\sigma_i} & |z - z_i| < \sigma_i, \\ 0 & \text{otherwise}, \end{cases}$$

(A2)

for the triangles. The widths of the bumps’ $\sigma_i$ can in general be redshift dependent, enabling us to differently sample different intervals. However, throughout this work, we have taken them to be constant. Triangular bumps were used earlier in the principal component analysis of different reionization scenarios (Hu & Holder 2003; Mortonson & Hu 2008). In some circumstances, the sharp edges in the triangles could cause numerical problems. Smoothed localized functions such as Gaussians and the $M_4$ splines introduced below therefore have a numerical advantage.

Instead of Gaussian and triangular bumps, one can also adopt an approach similar to that used in SPH, and think of the basis functions as window functions (or kernels) used to interpolate the properties of particles to any point in the medium. For us, the particles would be the spline knots (e.g., see De Boor 2001) at the specific $z_i$ with the associated magnitude $y_i$. There is a smoothing length $h$ associated with the kernel over which the properties of the particles are smoothed. Another commonly used kernel (other than the Gaussian considered above) is the cubic $M_4$ spline (e.g., Monaghan 2005), defined by

$$\varphi_i(z) \propto M_4(|z - z_i|) = \begin{cases} 1 - \frac{(2 - q)^3}{6} - 4(1 - q)^3 & 0 \leq q \leq 1; \\ \frac{1}{6}(2 - q)^3 & 1 \leq q \leq 2; \\ 0 & q > 2; \end{cases}$$

(A3)

where $q = |z - z_i|/h$. Whereas the Gaussian kernel has non-zero contributions from every redshift (though the range is usually truncated beyond about $3\sigma$), the cubic spline is compact, reaching zero for particles beyond $2h$.

As mentioned above, in this work the width of the bumps of these mode functions is chosen to be independent of redshift. We choose $\sigma_i = \delta z/2$ for Gaussian and triangular bumps (Equations (A1) and (A2)) and $h = 1.5\delta z$ for $M_4$ splines (Equation (A3)). In all cases, $\delta z = \Delta z/(N + 1)$ is the spacing between the centers of adjacent bumps, where $\Delta z$ is the redshift range of interest and $N$ is the number of basis functions used.

As basis functions, it is more convenient if the set of $\varphi_i$ is an orthogonal set. For this, there should be no overlap between different bumps. On the other hand, there is no way to cover the whole redshift range—a necessary condition for completeness—with a finite number of non-overlapping bumps. However, depending on the problem of interest, the width and separation of the (overlapping) bumps can be properly chosen to ensure all points in the redshift interval have been covered, while at the same time the orthogonality is not strongly violated.

APPENDIX B
NON-LOCALIZED BASIS FUNCTION

The most commonly used set of non-localized basis functions is the Fourier series:

$$u_i(z) \propto \cos(i\pi y) i = 0, 1, 2, \ldots$$

(B1a)

$$u_i(z) \propto \sin(i\pi y) i = 1, 2, \ldots$$

(B1b)

$$y = \frac{z - z_{\text{mid}}}{\Delta z/2}.$$  \hspace{1cm} (B1c)

where $\Delta z$ and $z_{\text{mid}}$ are the width and central point in the redshift range of interest. Thus, we have $|y| \leq 1$ as is required for Fourier expansion.

Alternatively, we can use Chebyshev polynomials of the first kind, $T_i$, to form the basis. These modes are constructed using the recursion formula:

$$T_{i+1}(y) = 2y T_i(y) - T_{i-1}(y).$$

with initial conditions $T_0(y) = 1$ and $T_1(y) = y$, and $y$ is given by Equation (B1c). Chebyshev polynomials of the

22
where in the language of Bayesian analysis, $L_{q_1}$ are orthogonal with respect to the weight function $w(y) = 1/\sqrt{1 - y^2}$.

These non-localized mode functions, unlike the localized case (with finite number of basis functions), do not suffer from non-orthogonality. However, both localized and non-localized sets of basis functions can in practice be considered complete if sufficiently many functions are taken into account.

APPENDIX C

EIGENMODES WITH FIXED AND VARYING BACKGROUND COSMOLOGY

For a model described by $N$ possibly correlated parameters, there is in general an orthogonal transformation which linearly couples between modes of different scales can be ignored, in the limit of full sky observation, or in cut-sky cases where can be written as

$$\text{Fisher matrix} = \sum_{\ell \geq 2} \frac{1}{2} \left( C^{-1} \frac{\partial C}{\partial q_i} C^{-1} \frac{\partial C}{\partial q_j} - \frac{\partial C}{\partial q_i} \frac{\partial C}{\partial q_j} \right).$$

In the case of fixed standard parameters, the Fisher matrix reduces to

$$F_{ij} = \frac{1}{2} \int \left( C^{-1} \frac{\partial C}{\partial q_i} C^{-1} \frac{\partial C}{\partial q_j} \right) dz.$$ 

In the standard CMB analysis with Gaussian signal and noise, we have $\mathcal{L} = \exp(-\Delta T^2/\Delta^2)/\sqrt{2\pi\Delta T}$. Here $\Delta$ represents the temperature and polarization maps including CMB signal as well as instrumental noise and $C = \langle \Delta T^2 \rangle$ is the theoretical pixel–pixel covariance matrix. With this likelihood function, the Fisher matrix simplifies to

$$F_{ij} = \frac{1}{2} \int \left( C^{-1} \frac{\partial C}{\partial q_i} C^{-1} \frac{\partial C}{\partial q_j} \right) dz.$$ 

The eigendecomposition of $F_{pp}$ then only requires the inversion of the well-behaved standard parameter block. It is straightforward to directly check that $(\mathbf{F}^{-1})_{pp}$ properly describes the marginal likelihood of the perturbation parameters:

$$\mathcal{L}(p|d) \propto e^{-p^T \mathbf{F}_{pp}^{-1} p}/2 \int e^{-s^T \mathbf{F}_{ss}^{-1} s}/2 e^{-\mathbf{p}^T \mathbf{F}_{ps}^T s} ds$$

Here $p$ and $s$ are the arrays of the perturbation and standard parameters.

The eigenmodes we are looking for can now be constructed using the eigenvectors of the inverse of the $pp$ block of the Fisher matrix $(\mathbf{F}_{pp})^{-1} = \mathbf{S}^{-1} \mathbf{S}^T$, where the columns of $\mathbf{S}$ are the eigenvectors of $(\mathbf{F}_{pp})^{-1}$ with their corresponding (non-negative) eigenvalues on the diagonal of the real diagonal matrix $\mathbf{F}^{-1}$. The eigenmodes we are looking for can now be constructed using these eigenvectors of the Fisher matrix and the basis functions we started with:

$$E_k(z) = \sum_{i=1}^{N} S_{ik} \phi_i(z).$$

If the $\phi_i$ happen to be orthonormal, then we have

$$\int_{z_{\text{min}}}^{z_{\text{max}}} E_k(z)E_l(z)w(z)dz = \delta_{kl}.$$
Figure 24. Three most constrained XeMs for three different fiducial models. The default model corresponds to the SRS and the effect of gravitational lensing on the CMB anisotropies has been included. One model corresponds to a recombination history with a different CMB temperature and in the other model lensing is not included. For the case of the two different CMB temperatures, the major difference is the shift in the eigenmodes associated with the shift in the fiducial Xe and visibility functions (see Figure 1).

(A color version of this figure is available in the online journal.)

Figure 25. Three most constrained XeMs with and without polarization and with $\ell_{\text{max}} = 2000$ and 3500.

(A color version of this figure is available in the online journal.)

Here $w(z)$ is the weight function with respect to which $\phi_i$ are orthonormal. Since Equation (C4) is not necessarily fulfilled, we force the $E_k(z)$ to be normalized to unity (as a matter of convenience), which is equivalent to a renormalization of the eigenvectors of $F$. Although in general this could change the rank ordering of the modes, in our case a reordering was not required for the modes included in the analysis. Now, instead of the original $\phi_i$, the set of the eigenmodes can serve as basis functions for the expansion of perturbations (compare with Equation (3)):

$$\delta u(z) = \sum_{k=1}^{N} \mu_k E_k(z) + r(z).$$

In Section 3.7, we discuss two examples of perturbation reconstruction with different numbers of eigenmodes taken into account (Figure 12). We demonstrate how well these eigenmodes serve as basis functions and also which features of the original perturbations are restored (or lost) if only a subset of the eigenmodes are used in the reconstruction process.

The square root of the eigenvalues of the inverse of the Fisher matrix can be used to forecast the error bars of the eigenmodes, i.e., $f_{ij} = \sigma_i^{-2} \delta_{ij}$, assuming that the probability distribution of the parameters is multivariate Gaussian close to the maximum. For non-Gaussian likelihoods, the $\sigma_i$ give the lower bound for the errors. In the rest of this paper we use the term error for the $\sigma_i$, as the Gaussianity of the likelihood function close to its maximum is usually a good assumption. If the modes are sorted in descending order of eigenvalues, the first few (with smallest $\sigma_i$) are the most constrainable. Thus, the constrainable part of the perturbations to the ionization history can be described by the eigenmodes which have reasonably small uncertainties (i.e., high eigenvalues), while the rest is practically unconstrainable by the data set under consideration.

APPENDIX D

FIDUCIAL MODEL AND DATA SET DEPENDENCE

The eigenmodes are by construction fiducial model dependent. In principle, the observables (such as $C_l$) for different fiducial models respond differently to the same perturbations depending on the strength of the signals, at different redshifts, from the unperturbed fiducial model.
As an example, in Figure 24 we compare the eigenmodes for three fiducial $X_e$ histories. Two of the models have different CMB temperatures and in the third one lensing has not been included. In the first two, the different $T_{\text{CMB}}$ lead to different fiducial $X_e$. Here, the main difference in the eigenmodes is their shift toward lower $z$ for the case with higher CMB temperature. This is consistent with the delayed recombination shown in Figure 1, remembering that XeMs are primarily localized around the maximum of visibility where the $C_\ell$ are most sensitive.

For the latter case with no lensing, although $X_e$ and the physics around recombination have not changed, there are still slight changes in some of the XeMs as seen in Figure 24.

We also checked the robustness of the eigenmodes against changes in the fiducial value of other parameters and the assumed reionization scenario. We tried a different value for $\Omega_b$, as the parameter most strongly affecting the ionization fraction, $1\sigma$ away from its fiducial value. For the late reionization we tried an extended reionization scenario (i.e., $X_e = 1$ for $z \lesssim 6$, $X_e = 0.5$ for $6 < z \lesssim 30$, and $X_e = 0$ elsewhere) radically different from our sharp fiducial reionization model (the default in CAMB). For both of these tests the first six eigenmodes were found to be the same as our main eigenmodes (Figure 5) with tiny differences in the fifth and sixth modes for the latter case.

This implies that, although the eigenmodes are fiducial model dependent, the constrainable ones are not practically sensitive to changes in the fiducial model or its parameters in the limits currently allowed by the data for the standard model of cosmology. That is because small changes in the fiducial parameters and the corresponding small changes in the ionization history only affect the XeMs at second order. Here by small we mean changes that lead to (smaller than or) the same order of magnitude signal in the simulated data as the (few best) XeMs. The higher XeMs with larger uncertainties are more affected by the same changes in the fiducial parameters, as these changes are no longer considered small relative to these poorly constrained XeMs. This non-sensitivity of the best modes to the fiducial values of parameters does not contradict their significant correlation once the standard parameters are also allowed to vary, as we see in Section 4.1.

We also studied the dependence of the XeMs on some properties of the simulated CMB data sets used for their construction, such as different $\ell_{\text{max}}$ corresponding to the smallest scale information present in the data, and different experimental noise levels. The results for a CVL experiment up to $\ell = 2000$ in temperature and polarization and also a CVL experiment only sensitive to temperature (up to $\ell = 3500$) are shown in Figure 25. As a more experimentally motivated case, we calculated the XeMs for simulated Planck-like data16 (using 100, 143, and 217 GHz channels, with effective galaxy-cut sky coverage of 75%) and ACTPol-like data, including both wide and deep surveys (Niemack et al. 2010). As shown in Figure 26, there is a tiny shift in the first mode relative to the mode for an ideal experiment and the changes grow as we proceed to higher modes.

More significant than the small changes in the XeMs constructed with different assumptions about data are the forecasted error bars in different cases (see Table 1). By removing the temperature at high $\ell$ or the polarization spectrum, the constraints on the amplitudes of the modes, determined from the eigenvalues of the Fisher matrix, become considerably larger. All these errors are calculated with the standard parameters fixed. However, the considered cases illustrate the general behavior of the method. Taking into account the correlation between the perturbations and the standard cosmic parameters leads to relatively higher error bars, depending on the data set used (Table 2).

REFERENCES

Ali-Haimoud, Y., & Hirata, C. M. 2010, Phys. Rev. D, 82, 063521
Ali-Haimoud, Y., & Hirata, C. M. 2011, Phys. Rev. D, 83, 043513
Bond, J. R. 1996, in Cosmology and Large Scale Structure, ed. R. Schaeffer, J. Silk, M. Spiro, & J. Zinn-Justin (Amsterdam: Elsevier), 469
Bond, J. R., Efstathiou, G., & Tegmark, M. 1997, MNRAS, 291, L33
Chen, X., & Kamionkowski, M. 2004, Phys. Rev. D, 70, 043502
Chluba, J., 2010, MNRAS, 402, 1195
Chluba, J., & Sunyaev, R. A. 2006a, A&A, 458, L29
Chluba, J., & Sunyaev, R. A. 2006b, A&A, 466, 39
Chluba, J., & Sunyaev, R. A. 2008, A&A, 480, 629
Chluba, J., & Sunyaev, R. A. 2009, A&A, 501, 29
Chluba, J., & Sunyaev, R. A. 2010, MNRAS, 402, 1221
Chluba, J., & Thomas, R. M. 2011, MNRAS, 412, 748
Das, S., Marriage, T. A., Ade, P. A. R., et al. 2011, ApJ, 729, 62
De Boor, C. 2001, A Practical Guide to Splines (Berlin: Springer)
Dubrovich, V. K., & Grachev, S. I. 2005, Astron. Lett., 31, 359
Dunkley, J., Hlozek, R., Sievers, J., et al. 2011, ApJ, 739, 52
Farhang, M., Bond, J. R., Doré, O., & Barth Netzerfield, C., 2011, arXiv:1108.2043
Fendt, W. A., Chluba, J., Rubio-Martin, J. A., & Wandelt, B. D. 2009, ApJS, 181, 627
Finkbeiner, D. P., Galli, S., Lin, T., & Slatyer, T. R. 2012, Phys. Rev. D, 85, 043522

16 http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI/200511_V2.pdf
