

Joint QSO - CMB constraints on reionization history

S Mitra

Harish-Chandra Research Institute, Chhatnag Road, Jhusi, Allahabad 211019, India

E-mail: smitra@hri.res.in

Abstract. We have tried to give an overview of model-independent semi-analytical approach to study the observational constraints on reionization. We have implemented and investigated a method to do a detailed statistical analysis using principal component analysis (PCA) technique. We have also discussed different observations related to reionization and shown how to use PCA for constraining the reionization history. Using Markov Chain Monte Carlo methods, we have found that all the quantities related to reionization can be severely constrained at $z < 6$, whereas a broad range of reionization histories at $z > 6$ are still permitted by the current data sets. We have shown that with the forthcoming PLANCK data on large-scale polarization, the $z > 6$ constraints will be improved considerably.

1. Introduction

In the past few years, the understanding of reionization process has become increasingly sophisticated in both the observational and theoretical communities, thanks to the availability of good quality data related to reionization. Mainly, the observations by the Wilkinson Microwave Anisotropy Probe (WMAP) satellite of cosmic microwave background (CMB) and highest redshift quasars put very tight constraints on the reionization history of the universe. The physical processes relevant to reionization are so complex that neither the analytical nor the numerical simulations alone can capture the overall picture. That is why, it is often studied using semi-analytical models of reionization, with limited computational resources. The semi-analytical model used in this work is based on Choudhury and Ferrara [1, 2]. This model treats the inter-galactic medium (IGM) as multi-phase medium, and takes into account the inhomogeneities in the IGM. The sources of ionizing radiation taken, here, are stars and quasars. The model computes *radiative feedback* (suppressing star formation in low-mass haloes using a Jeans mass prescription) self-consistently from the evolution of the thermal properties of the IGM. The major uncertainty in modelling any semi-analytical reionization scenario is to model the parameter N_{ion} , the number of photons entering the IGM per baryon in collapsed objects.

2. Principal component analysis

Most likely, N_{ion} is a function of halo mass and the redshift, but the dependences are not well understood, and that is why it is usually taken to be a constant. A method which is ideally suited to tackle this problem is to use the principal component analysis (PCA); this is a technique to compute the most meaningful basis to re-express the unknown parameter set, and the hope is that this new basis will reveal hidden detailed statistical structure [3, 4].

In this paper, we have assumed $N_{\text{ion}}(z)$ to be an unknown function of z and decompose it into principal components. We have included only one stellar population and hence, there



should not be any explicit chemical feedback in our model. A detailed likelihood analysis is done using three different data sets: The photoionization rates Γ_{PI} , obtained using Ly- α forest Gunn-Peterson optical depth observations and a large set of hydrodynamical simulations, the redshift distribution of LLS dN_{LL}/dz at $0.36 < z < 6$, and the angular power spectra C_l for TT, TE and EE modes using WMAP-7 and forecasted PLANCK data. The moment we have included the C_l 's (TT+TE+EE) in our analysis, we have realized that parameters related to reionization may have strong degeneracies with (some of) the cosmological parameters, mainly with the normalization of the matter power spectrum σ_8 and the slope n_s , so we have carried out our analysis by varying only these two parameters (in addition to the parameters related to reionization model) and keeping all the other cosmological parameters fixed to their best-fit WMAP-7 value.

2.1. Formalism

We have represented the unknown function $N_{\text{ion}}(z)$ by a set of n_{bin} discrete free parameters with redshift bins. Then we have constructed the Fisher matrix:

$$F_{ij} = \sum_{\alpha=1}^{n_{\text{obs}}} \frac{1}{\sigma_{\alpha}^2} \frac{\partial \mathcal{G}_{\alpha}^{\text{th}}}{\partial N_{\text{ion}}^{\text{fid}}(z_i)} \frac{\partial \mathcal{G}_{\alpha}^{\text{th}}}{\partial N_{\text{ion}}^{\text{fid}}(z_j)}, \quad (1)$$

where $\mathcal{G}_{\alpha}$, $\alpha = 1, 2, \dots, n_{\text{obs}}$ represent the observational data points, σ_{α} are their corresponding error, $\mathcal{G}_{\alpha}^{\text{th}}$ is theoretical value of $\mathcal{G}_{\alpha}$ and $N_{\text{ion}}^{\text{fid}}$ is the fiducial model, which is, in principle, close to the underlying “true” model. Here, we have taken the fiducial model $N_{\text{ion}}^{\text{fid}}$ to be the one, which matches the current data points up to an acceptable accuracy, and also which is characterized by a higher N_{ion} at higher redshifts. Once the Fisher matrix is constructed, we can determine its eigenvalues and corresponding eigenvectors. Because of the ortho-normality and completeness of the eigenfunctions, we can expand the deviation of N_{ion} from its fiducial model as $\delta N_i = N_{\text{ion}}(z_i) - N_{\text{ion}}^{\text{fid}}(z_i) = \delta N_i = \sum_{k=1}^{n_{\text{bin}}} m_k S_k(z_i)$, where $S_k(z_i)$ are the principal components of $N_{\text{ion}}(z_i)$, and m_k are the expansion coefficients. The advantage is that, unlike $N_{\text{ion}}(z_i)$, the coefficients m_k are uncorrelated with variances. It can be shown that the largest eigenvalues correspond to minimum variance and vice versa. Hence, most of the information relevant for the observed data points is contained in the first few modes with larger eigenvalues. We can then reconstruct the function δN_i using only the first $M \leq n_{\text{bin}}$ modes. So, the important step in this analysis is to decide on how many modes M to be used.

3. Results: WMAP-7 vs PLANCK

3.1. The principal components of $N_{\text{ion}}(z)$

We have found that the components of the Fisher matrix vanish for $z < 2$, because there are no data points considered at these redshifts. For $2 < z < 6$, the values of F_{ij} are considerably higher, because it is determined by the sensitivity of Γ_{PI} and dN_{LL}/dz on $N_{\text{ion}}(z)$. On the other hand, the information at $z > 6$ is determined by the sensitivity of C_l^{EE} on N_{ion} , which is relatively weak. We have found that F_{ij} is negligible for $z > 14$, and this is expected because the collapsed fraction of haloes is negligible at those redshifts, and hence, there exist no free electrons to contribute to electron scattering optical depth τ_{el} or C_l^{EE} [5, 6]. In Figure 1, we have shown the inverse of the first few eigenvalues (λ_i), i.e., the variances of the corresponding modes. For modes $i > 8$, the eigenvalues are almost zero, and the variances are extremely large. This implies that the errors on N_{ion} would increase dramatically if we include modes $i > 8$. The first 8 eigenmodes (i.e., those which have the lowest variances) are plotted in Figure 2. We have found that all the eigenmodes tend to vanish at $z < 2$ and $z > 15$, which is obvious because of F_{ij} being negligible at these redshifts. We can see a number of spikes and troughs in the first four modes, whose positions correspond to the presence of data points for Γ_{PI} and dN_{LL}/dz at $2 < z < 6$.

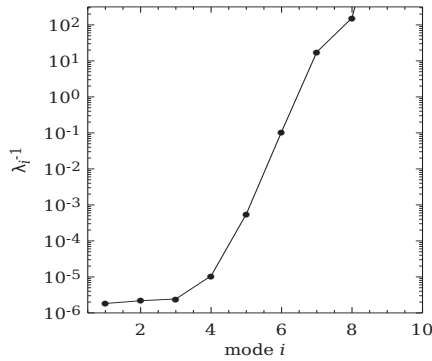


Figure 1. The inverse of eigenvalues of the Fisher matrix F_{ij} , which essentially measures the variance on the corresponding coefficient.

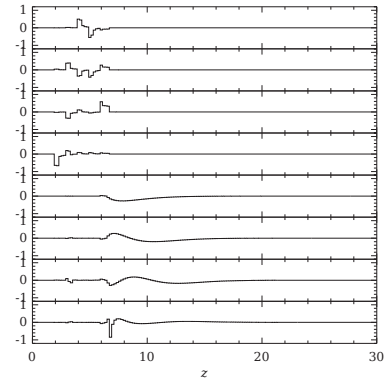


Figure 2. The first 8 eigenmodes of the Fisher matrix.

The last four modes contain the information about the sensitivity of C_l^{EE} . The modes (> 8) with smaller eigenvalues, i.e., large variances introduce huge uncertainties in the determination of N_{ion} , and hence, do not contain any meaningful information about the reionization history.

3.2. Choice of the number of modes

The next step in our analysis is to decide on how many modes M to use. A sophisticated and model-independent prescription is to use Akaike information criterion ($\text{AIC} = \chi_{\text{min}}^2 + 2M$), where smaller values of AIC are assumed to imply a more favoured model. Note that there is no reason to select one particular reconstruction, the minimum of AIC can be accompanied by an increased chance of getting the reconstructed parameters wrong. One successful strategy for this purpose is to select different M , which are near the minimum value of AIC and amalgamate them equally at the Monte Carlo stage when we compute the errors [7]. In this way, we can reduce the inherent bias, which exists in any particular choice of M . We have examined that, in our case, the family of different M reconstructions, starting from $M = 2$, which satisfy $\text{AIC} < \text{AIC}_{\text{min}} + \kappa$, where $\kappa = 10$ (which corresponds to $M = 8$) produces very solid results. For alternative data sets, the value of κ can be adjusted.

3.3. Markov Chain Monte Carlo constraints

The constraints on reionization are obtained by performing a Monte Carlo Markov Chain (MCMC) analysis over the parameter space of the optimum number of PCA amplitudes and other free parameters of the model. To avoid the confusion about the correct choice of number of modes, we have performed the MCMC analysis for PCA amplitudes taking from $M = 2$ to $M = 8$, all of which obey the AIC criterion.

It can be seen from the plot of $N_{\text{ion}}(z)$ (*top-left* panel of Figure 3) that such quantity must necessarily increase from its constant value at $z < 6$. This rules out the possibility of reionization with a single stellar population having non-evolving IMF and/or star-forming efficiency. From the plot of $\Gamma_{\text{PI}}(z)$ (*top-middle* panel) and dN_{LL}/dz (*top-right* panel), we have found that the mean model is consistent with the observational data at $z < 6$, as expected. The errors corresponding to 95% confidence limits are also smaller at this epoch. However, the model described by the mean values of the parameters shows a sharper and prominent peak around $z \sim 6.5$ for both cases. From the plot of the volume filling factor of ionized region $Q_{\text{HII}}(z)$ (*bottom-left* panel), we have seen that the growth of this quantity for the mean model is much faster than that of fiducial model at initial stages, though the completion of reionization takes

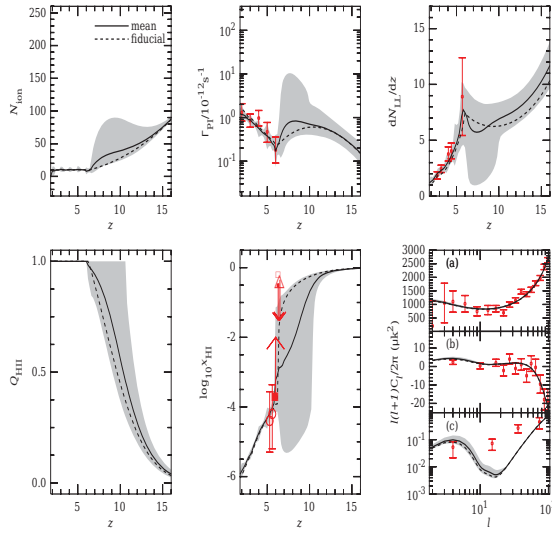


Figure 3. The marginalized posteriori distribution of various quantities related to reionization history obtained from the PCA using the AIC criterion with first 8 eigenmodes for WMAP-7 data.

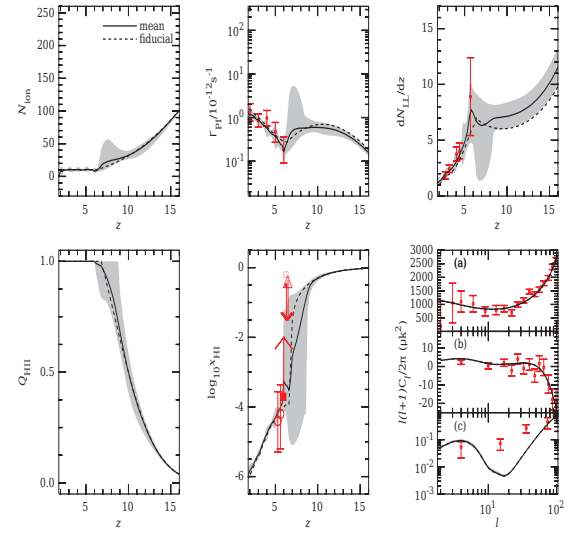


Figure 4. The marginalized posteriori distribution of various quantities related to reionization history obtained from the PCA using the AIC criterion with first 8 eigenmodes for PLANCK likelihood.

place only at $z \approx 6$. Similarly, the neutral hydrogen fraction $x_{\text{HI}}(z)$ (*bottom-middle* panel) decreases much faster than the fiducial one at $6 < z < 12$ and then smoothly matches the Ly- α forest data. Finally, we have shown the values of (a) C_l^{TT} , (b) C_l^{TE} and (c) C_l^{EE} for the mean model in the *bottom-right* panel of this figure, which is almost the same as the fiducial model. Also, here, we have got the value of τ_{el} higher than the current WMAP value. So, a wide range of reionization histories is still allowed by the data we have used.

To check how better the PLANCK data can constrain the reionization scenario, we have first generated the simulated PLANCK data of CMB power spectra for our fiducial model up to $l \leq 2000$ using the exact full-sky likelihood function at PLANCK-like sensitivity, and then repeat the same MCMC analysis as we have done for WMAP-7 data. In Figure 4, we have illustrated the recovery of the same quantities as mentioned in the earlier case. We have found that our main results are in quite reasonable agreement with those obtained from the WMAP data, except that all the 2σ (95 %) limits are reduced remarkably for all redshift range, especially the 2σ limits for Q_{HII} reduces significantly for this case. Thus, with the forthcoming PLANCK data on large scale polarization (ignoring the effect of foregrounds), the $z > 6$ constraints will be improved considerably. In particular, we have found that using the PLANCK data, the 2σ error on τ_{el} will be reduced to 0.009 and the uncertainties on the redshift at which reionization is 50% completed, i.e., $z(Q_{\text{HII}} = 0.5)$ and almost or 99% completed, $z(Q_{\text{HII}} = 0.99)$, would be ~ 1 and 3 (95% CL) respectively, which are much smaller than the case for current WMAP data. However, for more stringent constraints on reionization at $z > 6$, one has to rely on the data sets other than CMB.

Acknowledgements

The author would like to thank all his collaborators, who have contributed to the research work described in this article.

References

- [1] Choudhury T R and Ferrara A 2005 *MNRAS* **361** 577
- [2] Choudhury T R and Ferrara A 2006 *MNRAS* **371** L55
- [3] Mortonson M J and Hu W 2007 *Ap. J.* **657** 1
- [4] Guha Sarkar T, Mitra S, Majumdar S and Choudhury T R 2012 *MNRAS* **421** 3570
- [5] Mitra S, Choudhury T R and Ferrara A 2011 *MNRAS* **413** 1569
- [6] Mitra S, Choudhury T R and Ferrara A 2012 *MNRAS* **419** 1480
- [7] Clarkson C and Zunckel C 2010 *Phy. Rev. Lett.* **104** 211301