High redshift intergalactic medium: Probes and physical models

Shiv K. Sethi

Raman Research Institute, Bangalore 560 080, India

Recent years have seen major advances in understanding the state of the intergalactic medium (IGM) at high redshift. Some aspects of this understanding are reviewed here. In particular, we discuss: (1) Different probes of IGM like Gunn–Peterson test, CMBR anisotropies, and neutral hydrogen emission from reionization, and (2) some models of reionization of the universe.

I. Introduction

One of the outstanding issues in cosmology is to understand the reionization of the universe. Following recombination of primordial plasma at redshift $z \simeq 1000$, the universe is mostly neutral with an ionization fraction $\simeq 10^{-4}$ (see e.g. refs 1–3). The Jean's mass at recombination is $\simeq 10^6\, M_\odot$. At $z \lesssim 100$, the plasma thermally decouples from CMBR and its temperature decreases adiabatically, $T_m \simeq 1/a^2$, which leads to a further decrease in Jean's mass. During this 'dark and cold age' the density perturbations at scales above the Jean's scale can grow. Figure 1 shows the ionization and thermal history of the universe along with the evolution of Jeans' mass. This age comes to an end when the first structures can collapse and the light from these objects can reionize and reheat the universe (for a recent review see ref. 4 and references therein). Therefore the epoch of re-ionization holds important clues about the way first structures formed and can potentially distinguish between different models of structure formation. In the standard ΛCDM models the first structures to collapse would be just above the Jean's length (see e.g. ref. 2). Another crucial question in this regard is whether these structures could cool fast enough to form stars. Many of these issues will be discussed in this review.

Different ongoing and potential probes of intergalactic medium can reveal the nature of the re-ionization of the universe. One of the most important and the oldest is the Gunn–Peterson (GP) test (see e.g. ref. 1 and references therein), which is very sensitive to the neutral fraction in the intergalactic medium. CMBR anisotropy measurements are another powerful and complementary probe, as they are sensitive to the ionized component of the intergalactic medium (see refs 5, 6 and references therein). In future, it might be possible to directly observe the first sources that re-ionize the universe. In addition, the transition from neutral to ionized universe might also be detected in neutral hydrogen emission (see e.g. refs 7–9).

To sum up the observational status: Recent detection of temperature-polarization cross-correlation in CMBR suggests that the redshift of reionization[10] $z_{\rm reion} \simeq 17 \pm 4$. GP probes suggest that the universe is highly ionized up to $z \simeq 5$, but might be making a transition from highly ionized to neutral[11–13] for $5 \lesssim z \lesssim 6$. These two observations together throw open the interesting possibility that the universe went through two phases of re-ionization.

This article is divided into two parts. In the first part, probes that give a clue about the reionization epoch will be discussed. In the second part, we will discuss the nature of ionizing sources. Throughout this review we use the currently-favoured FRW model: spatially flat with $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$ (refs 14–16) with $\Omega_b h^2 = 0.02$ (refs 14, 17) and $h = 0.7$ (ref. 18).

II. Probes of ionization at high redshifts

Gunn–Peterson effect

High redshift sources should show absorption at frequencies close to Lyman-α (1216 Å) line owing to scattering from

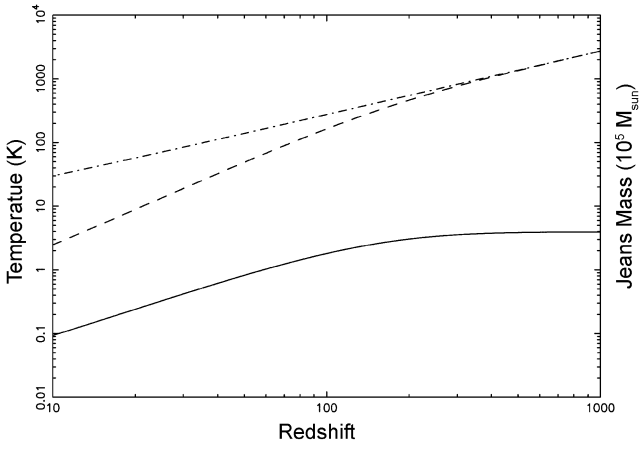

Figure 1. The temperatures of CMBR and matter are plotted alongside the evolution of Jeans mass. The solid curve is Jeans mass in units of $10^5\, M_\odot$. The dashed and dot-dashed curves correspond to matter and CMBR temperature (in Kelvin), respectively.

e-mail: sethi@rri.res.in

IGM neutral hydrogen (HI). This test applied to high redshift quasars which generally have a strong Lyman-α emission line means that the blueward side of the Lyman-α line should show strong absorption as compared to the redward side (GP test). The optical depth to the Lyman-α scattering from HI can be calculated (see e.g. ref. 1):

$$\tau_{\rm GP}(n_0) \simeq 4 \times 10^5 x_{\rm HI} \left(\frac{h}{0.7}\right)\left(\frac{\Omega_b}{0.045}\right)\left(\frac{0.3}{\Omega_m}\right)^{1/2}\left(\frac{1+z}{6}\right)^{1.5}. \quad (1)$$

Here $x_{\rm HI} \equiv n_{\rm HI}/n_{\rm H}$ ($n_{\rm H} \simeq 0.92\ n_b$) is the neutral fraction of hydrogen and $n = n_\alpha/(1 + z)$. Two points worth noting in eq. (1) are: (1) Owing to the resonant nature of Lyman-α scattering, optical depth at observed frequency n_0 gives direct information about the neutral fraction at redshift $(1 + z) = n_\alpha/n_0$, and (2) more importantly, this test is extremely sensitive to the neutral fraction of hydrogen, even a neutral fraction as small as a part in hundred thousand can fully absorb light shortward of Lyman-α in the quasar spectrum.

Since 1960s when the first high redshift quasars whose Lyman-α emission could be detected from ground-based telescopes were discovered, GP test has been applied to study the ionization of the universe at high redshifts. Till 2000, none of the observed quasars up to a redshift $\simeq 5$ showed any GP absorption, which means the universe is ionized to better than a part in a million up to $z \simeq 5$. The discovery of several quasars at redshifts above six by SDSS survey made it possible to apply GP tests at even higher redshifts. GP absorption has been detected[11–13] in several quasars at redshifts $z \gtrsim 5.7$. Observations suggest that the neutral fraction increases rapidly between $z \simeq 5.5$ and $z \simeq 6.2$ (see e.g. ref. 11). These observations might mean that the universe is becoming ionized owing to the formation of the first structures at $z \simeq 6$. However a straightforward interpretation of these results is not easy. Even for the best quasar spectrum and using the Lyman-β line, which has a smaller oscillator strength, the GP optical depth[19] $\tau_{\rm GP} \gtrsim 25$. Using eq. (1), this implies that $x_{\rm HI} \gtrsim 10^{-4}$, i.e. the universe can be almost fully ionized. Use of semi-analytic models which take into account the clumpiness of the IGM give more stringent constraints on the neutral fraction and give[11] $x_{\rm HI} \gtrsim 10^{-3}$ at $z \simeq 6$. Even though the observations are unable to conclude that the universe made a transition from fully ionized to almost fully neutral between redshift of five and six, the ionized fraction certainly evolves very rapidly in this redshift range, and this can have important implications for the models of structure formation.

CMBR anisotropies

CMBR anisotropies provide a complementary approach to the ionization history of the universe as compared to the GP test, as they are sensitive to the ionized component of the universe. The physics of CMBR temperature anisotropies at the last scattering surface and various data analysis issue are covered elsewhere in this volume (Subramanian, this volume). Here we shall discuss the implications of reionization on the CMBR temperature anisotropies. To highlight the effect of reionization on the polarization anisotropies we discuss more fully the CMBR polarization anisotropies.

The universe recombines at $z \simeq 1000$. Following recombination the ionized fraction ($x_e \equiv n_e/n_{\rm H}$) in the universe is $\simeq 10^{-4}$ for $z \simeq 100$ (see e.g. ref. 1). The mean-free path of the CMBR photons to Thompson scattering exceeds the local Hubble radius in the post-recombination era and therefore the universe is 'transparent' to the CMBR photons. Following reionization, the ionized fraction might reach nearly unity and a small fraction of CMBR photons might re-scatter again. An important quantity in studying CMBR quantity is visibility function which is the normalized probability that the photon scattered in a range z and $z + {\rm d}z$. It is defined as: $V(h_0, h) = \dot{\tau} \exp(-\tau)$, here ${\rm d}h = {\rm d}t/a$ is the conformal time and $\tau = \int_{h_0}^{0} x_e n_{\rm H} a \sigma_t c {\rm d}h$. In Figure 2 we show the visibility function for two models. For the model with early reionization, the visibility function get important contribution from redshift of reionization. One can define the optical depth to the reionization surface: $\tau_{\rm reion} = \int_0^{z_{\rm reion}} x_e n_{\rm H} \sigma_t c {\rm d}t$; this is the fraction of photons that re-scattered in the reionized universe for $\tau_{\rm reion} < 1$. For $\tau_{\rm reion} \simeq 6$, $\tau_{\rm reion} \simeq 0.05$ and $\tau_{\rm reion} \propto (1 + z_{\rm reion})^{3/2}$).

The temperature and polarization anisotropies can be computed by solving the Boltzmann equations for the photon distribution function. Equations appropriate for studying the effect of reionization for scalar perturbations, for a given wavenumber **k** and line of sight **n**, are (see e.g. ref. 5, Zalddariaga[6,20]:

$$\dot{\Delta}_T + ik\mu\Delta_T \simeq \dot{\tau}(\mu v - \Delta_T)$$

$$\dot{\Delta}_P + ik\mu\Delta_P = \dot{\tau}(\Pi(\mu)[\Delta_{T^2} + \Delta_{P2} - \Delta_{P0}] - \Delta_P). \quad (1)$$

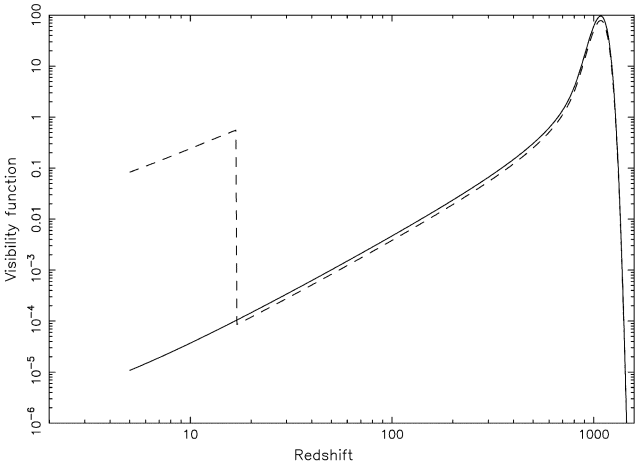

Figure 2. Visibility function, defined as ${\rm d}t/{\rm d}h \exp(-\tau) H_0^{-1}$, is plotted for different models. The solid and the dashed curves are for the standard recombination and a model in which the universe reionizes at $z = 17$, respectively.

Here $m = \mathbf{k}.\mathbf{n}$, $\dot{\Delta}_T \equiv \partial\Delta_T/\partial h$, $\Delta_T \equiv \Delta T/T(k, m, h)$. The polarization anisotropies $\Delta_P \equiv \Delta P/T(k, m, h)$ are in the plane perpendicular to the **k** vector, which is taken to be parallel to the *z*-axis, to exploit the axial symmetry of the problem. In this case, only one Stokes parameter Q is non-zero and $\Delta P = QT$. Also $\dot{t} = n_b x_e s_T a$, $\Pi(m) = 0.5(1–P_2(m))$. v is the electron velocity; we also assume curl-free velocity fields which allow us to express: $\mathbf{v}.\mathbf{n} = \mathbf{v}.\mathbf{k}$. The angular moments of temperature and polarization anisotropies are: $\Delta_\ell = \frac{1}{2}\int_{-1}^{1} dm P_\ell(m)\Delta(k, m)$. Even though the temperature equation is only appropriate for studying the effect of reionization on anisotropies generated at last scattering surface, the polarization equation is exact and can also be used to study the generation of perturbations at the last scattering surface.

To understand the essential physics of CMBR anisotropies in reionized models, we only solve for anisotropy evolution for one wavenumber **k**. The quantity measured by experiments is the two-point function of this quantity summed over all the wave-numbers (for details see articles by Subramanian, this volume).

We begin by studying the effect of reionization on the temperature anisotropies. Equation (2) can be solved to give:

$$\Delta_T(h) = \Delta_T(h_{\rm rec})\exp(ikm(h_{\rm rec} - h))\exp(-t(h_{\rm rec}, h)) + m\int dh' v(h') V(h, h')\exp(ikm(h' - h)) \quad (3)$$

The first term in the solution means the anisotropies generated at the last scattering surface are exponentially damped by reionization (the solution is only correct for scales smaller than the size of local horizon, i.e. $k \gtrsim h^{-1}(z_{\rm reion})$. It is not reflected in the solution owing to dropping a term $\propto \Delta_{T0}$ in the temperature equation (for details see ref. 21). For example, if the universe reionized at $z \simeq 50$ which gives $t_{\rm reion} \simeq 1$, i.e. all the photons from the last scattering surface are re-scattered following reionization, this means that all anisotropies at scales smaller than the angular scale correspond to $\ell \simeq h_0/h\ (z_{\rm reion}) \simeq 10$ are wiped out. As we shall see the second term in the solution does not contribute much to the anisotropies at small scales either. This means that for a reionization redshift $\simeq 50$ no anisotropies should be observed for $\ell \gtrsim 10$, which is in direct contradiction with observations (e.g. WMAP observations detect anisotropies for $\ell \simeq 600$). Therefore the redshift of reionization should be small enough such that only a small fraction of CMBR photons are re-scattered. To compute the second term in eq. (3), we can assume the visibility to be a normalized Gaussian with a width $\Delta h_{\rm reion}$, this gives

$$\Delta_T \propto m t_{\rm reion} v \exp[-(km\Delta h_{\rm reion}/2)^2]. \quad (4)$$

This shows that CMBR anisotropies generated during the epoch of reionization owing to Doppler scattering off electrons are suppressed for $k \gtrsim 1/\Delta h_{\rm reion}$. The generated signal as expected is $\propto t_{\rm reion}$ for realistic ionization histories $\Delta h_{\rm reion} \simeq h_{\rm reion}$. We show in Figure 3, the effect of reionization on the temperature anisotropies. The net effect of reionization on the CMBR temperature anisotropies can be summarized as: (a) anisotropies at small scales are suppressed exponentially as $\exp(-2t_{\rm reion})$, (b) anisotropies at scales corresponding to $\ell \lesssim 10$ escape this exponential damping and also new anisotropies are generated at these scales. As seen in Figure 3, reionization causes a relative decrease in the small scale anisotropies. The minimum error in detecting the angular power spectrum of CMBR anisotropies at any ℓ is $\Delta C_\ell \simeq \sqrt{2/(2\ell + 1)}C_\ell$ (cosmic variance) and for all CMBR anisotropy experiments the error on C_ℓ is dominated by the cosmic variance for $\ell \lesssim 300$. The reionization signal is very difficult to detect owing to uncertainty in the overall normalization of the CMBR anisotropies which is compounded by cosmic variance. From WMAP data all the information about the reionization comes from the polarization signal.

In second order in perturbation theory, reionization causes potentially detectable anisotropies for $\ell \gtrsim 1000$; for $\ell \gtrsim 2000$–3000 this signal can dominate the signal from primary anisotropies. The most important second order contribution to temperature anisotropies comes from Vishniac effect (for details see ref. 22, also see refs 5, 20, 21 for host of other second order effects).

From eq. (2), it can be seen that the source of polarization anisotropy is the quadrupole of the temperature anisotropy. The generation of temperature and polarization anisotropies at the last scattering surface is generally an involved process and eq. (2) has to be solved numerically (see e.g. ref. 23). However, several relevant assumptions allow one to get an analytic insight into the problem[24,25]. First simplification occurs because the photon–baryon plasma can be treated as tightly-coupled for most of the period during which the

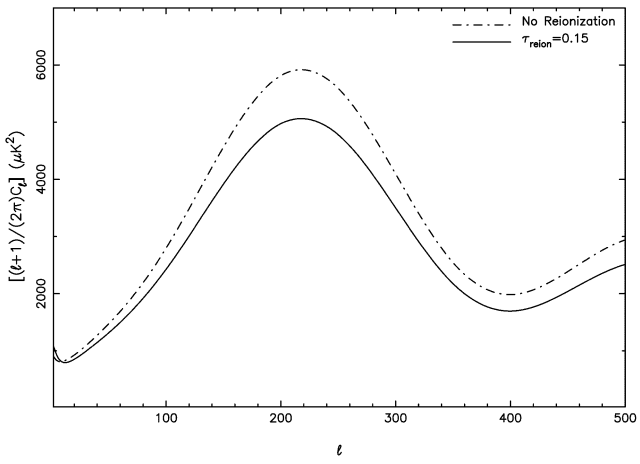

Figure 3. The effect of reionization CMBR temperature anisotropies is shown. The main effect of reionization on temperature anisotropies is to generate secondary signal for $\ell \lesssim 20$ and suppress primary anisotropies by a factor $\exp(-2t_{\rm reion})$ for larger ℓ.

recombination lasts. The tight coupling approximation is valid for scales corresponding to $k \lesssim l_f^{-1}$, where l_f is the (comoving) mean-free path of the photons for Thompson scattering; $l_f \simeq 1$ Mpc at the last scattering surface for a fully ionized universe, which means that tight coupling approximation is valid for most scales of interest. (The scale of interest for tight-coupling in not l_f but the scale of photon diffusion at recombination which for a fully ionized plasma is roughly 10 times larger than l_f; free-streaming of photons at recombination further increase this scale; for detailed discussions and implications of this for CMBR anisotropies see ref. 24. A second simplification, closely related to the first but physically distinct, occurs because the width of visibility function at the last scattering surface corresponds to scales $\lesssim 10$ Mpc, and hence for studying physical processes at much larger scales the recombination can be treated as instantaneous (for caveats see ref. 24). Our interest here is in scales that are super-horizon at the last scattering surface ($k^{-1} \gtrsim 100$ Mpc, comoving), corresponding to angular scales $\ell \lesssim 200$. Therefore we will use tightly coupled approximation and not discuss the effects of photon diffusion. Using these approximations, adiabatic initial conditions and also the fact that the ratio of baryon to photon energy density $r_b/r_g \simeq 25\ \Omega_b h^2 \ll 1$ at the epoch of decoupling for acceptable models of primordial nucleosynthesis, the temperature and polarization anisotropies generated at the last scattering surface are: (for details see refs 24, 25):

$$\Delta_T(k,\mathbf{n},h) = \frac{1}{3}\Phi(\mathbf{k},h_{\rm rec})\cos(kc_s h_{\rm rec})\exp[ikm(h_{\rm rec},-h)]$$

$$\Delta_P(k,\mathbf{n},h) = 0.17(1-m^2)kc_s\Delta h_{\rm rec}\Phi(\mathbf{k},h_{\rm rec})\sin(kc_s h_{\rm rec})\exp[ikm(h_{\rm rec}-h)]. \quad (5)$$

Here $\Phi(\mathbf{k}, h_{\rm rec})$ is the Fourier component of the Newtonian potential at the last scattering surface; $\Delta h_{\rm rec} \simeq 10$ Mpc is the comoving width of LSS; $c_s \simeq 1/\sqrt{3}$ is the sound velocity in the coupled photon–baryon fluid. These large scale solutions show that: (a) temperature and polarization anisotropies are correlated with each other, (b) the amplitude of polarization anisotropies is suppressed by a factor $kc_s\Delta h_{\rm rec}$; e.g. for $k \simeq 0.01$ Mpc^{-1} this factor is roughly 1/10. This suppression is owing to the fact that polarization anisotropies $\propto \Delta_{T2}$. In the strict tight coupling approximation, only the temperature monopole and dipole are non-zero in the comoving frame. A small quadrupole is generated owing to free streaming of photons before they scatter for the last time; and this quadrupole is suppressed with respect to the monopole and dipole, which constitute the primary sources of temperature anisotropies, by the factor $\simeq kc_s\Delta h_{\rm rec}$. As we shall see below, in the reionized models this suppression is absent. In Figure 4 we show the CMBR temperature and polarization anisotropies, generated at the last scattering surface.

In the reionized models, a fraction CMBR photons scatter again after the epoch of reionization. As argued above, this fraction is small as many of the features of the primary anisotropies have already been detected. We have already discussed the effect of reionization on temperature anisotropies and argued that the reionization signal is very difficult to discern from temperature anisotropies alone. Like temperature anisotropies, one of the effects of reionization would be to wipe out polarization anisotropies generated at the last scattering surface for scales $\lesssim h_{\rm reion}$. From eq. (2), the generation of new polarization anisotropies will be proportional to the value of Δ_{T2} at the epoch of reionization. Equation (2) can be simplified further by dropping terms of polarization monopole and quadrupole of polarization anisotropies in the RHS of the equation, as these terms are negligible compared to the temperature quadrupole (see below). Equation (2) then gives the following solution for the generation of polarization anisotropies following reionization:

$$\Delta_P = \int \mathrm{d}h'\Pi(m)\Delta_{T2}(h')V(h,h')\exp(ikm(h'-h)). \quad (6)$$

Most of the contribution to the integral will come from close to the reionization epoch (typically $\simeq \Delta h_{\rm reion} \simeq h_{\rm reion}$, the width of visibility function if it is approximated as a Gaussian as above). The amplitude of the contribution is proportional to the temperature quadrupole at the epoch of reionization. As argued above, the temperature quadrupole is suppressed at the last scattering surface owing to tight coupling approximation that holds at that epoch. Following recombination the photons free stream which allows to temperature quadrupole to increase. From eq. (5), we can get the temperature quadrupole at the reionization epoch from taking the angular moment of the equation (for details see. e.g. ref. 26):

$$\Delta_T^{(2)}(h_{\rm reion}) = \frac{1}{3}\Phi(k,h_{\rm rec})\cos(kc_s h_{\rm rec})\, j_2[k(h_{\rm reion}-h_{\rm rec})]. \quad (7)$$

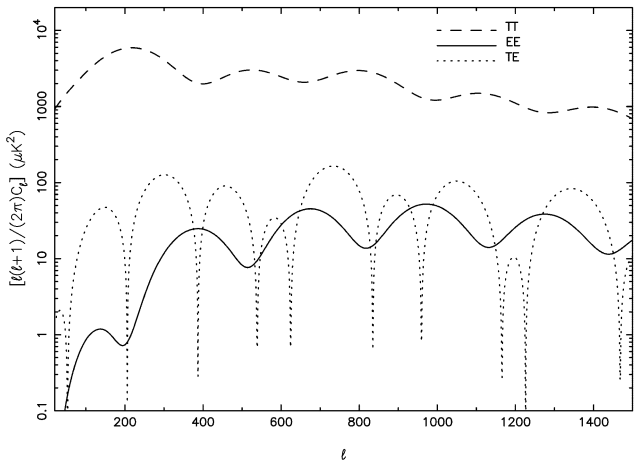

Figure 4. Primary CMBR temperature (TT) and polarization (EE) anisotropies and their cross-correlation (TE) are shown; the absolute value of the temperature-polarization cross-correlation is plotted.

Using the fact that the maximum value of $j_2(x) \simeq 0.3$, and comparing with the polarization anisotropies generated at the last scattering surface, eq. (5), we can see that $\Delta_T^{(2)}(h_{\rm reion})$ does not suffer the k dependent suppression, and is of the order of temperature anisotropies at the last scattering surface. At $k \ll 0.1$, it can be several orders of magnitude more than the polarization anisotropies generated at the last scattering surface. The polarization anisotropies at the reionization epoch peak at the characteristic scale $k \simeq 2/h_{\rm reion}$, which corresponds to an angular scale $l \simeq kh_0 \simeq 5$–10 for $z_{\rm reion} \simeq 15$–50.

In Figure 5 we show the effect of reionization on the temperature–polarization cross-correlation power spectrum; also shown are the observations of WMAP. Figure 5 shows that the enhancement of power at $\ell \leq 10$ cannot be explained within the framework of no reionization models. Also shown are the predictions of a model in which the universe reionized at $z \simeq 5.5$, which the GP observations discussed above might be suggesting and the best fit model to the WMAP observations. The best fit model requires $t_{\rm reion} \simeq 0.15$, which implies the epoch of reionization corresponds to $z_{\rm reion} \simeq 15$.

High redshift HI

Another possible probe of the reionization epoch is to observe the neutral component of hydrogen through the epoch of reionization. The neutral fraction of hydrogen changes from near unity to zero during the epoch of reionization. This change can potentially be observed using the hyperfine transition of the hydrogen atom at $n_* = 1420$ MHz. The quantity of interest here is the spin temperature of hydrogen defined as:

$$\frac{n_2}{n_1} = 3\exp(-T_*/T_s). \tag{8}$$

Here n_2 and n_1 are the populations of the hyperfine states. $T_* = hn_*/k = 0.06$ K. As the only radio source at the high redshift is CMBR, HI in hyperfine transition can be seen against the CMBR in emission or absorption. The observed quantity then is the deviation of CMBR from a black body at radio frequencies. The observed difference is:

$$\Delta T_{\rm CMBR} = -t_{\rm HI}(T_{\rm CMBR} - T_s). \tag{9}$$

Here $t_{\rm HI} = s_n N_{\rm HI} T_*/T_s$, with the HI column density $N_{\rm HI} = \int n_{\rm HI} {\rm d}\ell$; $s_n = c^2 A_{21} f_n/(4\pi n_*)$; $A_{21} \simeq 1.8 \times 10^{-15}$ s^{-1} and f_n is the line response function. The spin temperature is determined from detailed balancing between various processes that can alter the relative populations of the two levels[27,28]:

$$T_s = \frac{T_{\rm cmbr} + y_c T_K + y_a T_a}{1 + y_c + y_a}. \tag{10}$$

Here $y_c \propto n_{\rm H}$ and $y_a \propto n_a$, with n_a being the number density of Lyman-α photons correspond, respectively, to relative probabilities with which collisions between atoms and the presence of Lyman-α photons determine the level populations; T_K is the matter temperature. In the pre-reionization era, there are no Lyman-α photons, and therefore $y_a = 0$.

From eq. (9) it is clear that HI can be observed in either absorption or emission against CMBR depending on whether T_s is less than or exceeds $T_{\rm CMBR}$. At $z \simeq 1000$, $T_{\rm CMBR} = T_K$ and it follows from eq. (10) that $\Delta T_{\rm CMBR} = 0$. As seen in Figure 1, for $z \lesssim 100$, $T_K < T_{\rm CMBR}$ and therefore HI can be observed in absorption against CMBR if $y_c \gtrsim 1$. During reionization, y_a term can become important and owing to thermalization at frequencies close to Lyman-α, $T_a \simeq T_K$ (ref. 28). The temperature of the medium can also exceed $T_{\rm CMBR}$ from X-ray and Lyman-α heating (for details see refs 29, 30) which implies that $\Delta T_{\rm CMBR} > 0$. $\Delta T_{\rm CMBR}$ approaches zero as the reionization is completed. In Figure 6, we show $\Delta T_{\rm CMBR}$ as a function of the observed frequency for an ionization history in which the universe becomes

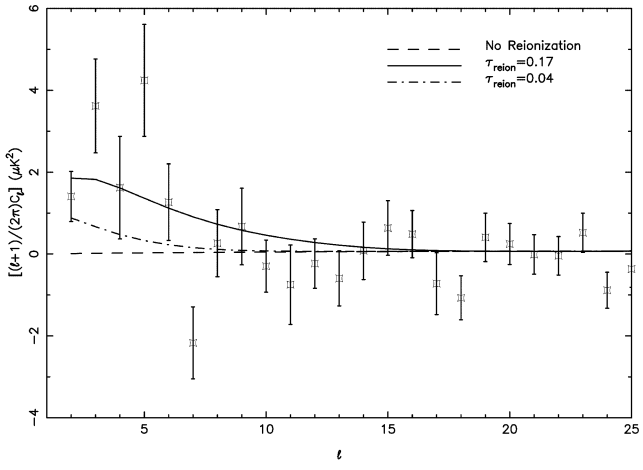

Figure 5. The effect of reionization on the temperature-polarization cross-correlation is shown. Also shown are the WMAP data points with one-sigma error bars[10].

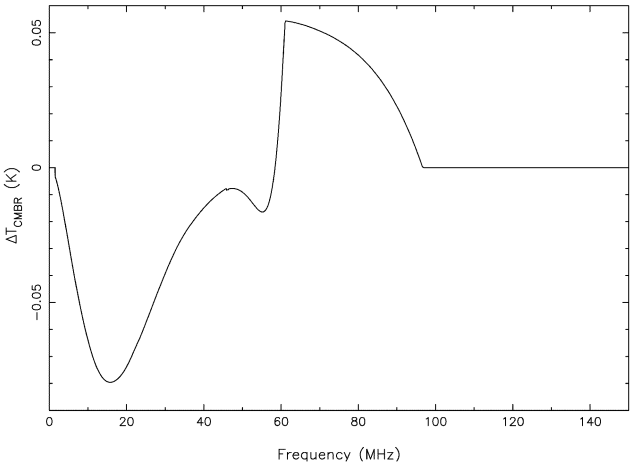

Figure 6. $\Delta T_{\rm CMBR}$ (eq. (9)) is shown as a function of the observed frequency for an ionization history in which the universe becomes fully ionized at $z = 15$.

fully ionized at $z = 15$ (more details in next section and ref. 30). If the universe is re-ionized at $z \simeq 15$, then $\Delta T_{\rm CMBR} \simeq 0.05$ K at frequencies $n \simeq 50$–80 MHz. In addition, there is a signal from pre-reionization epoch with $\Delta T_{\rm CMBR} \simeq -0.05$ K at $n \simeq 30$ MHz. In addition to the average signal, there will be fluctuations in the temperature difference owing to fluctuations in HI density from primordial density perturbations and also from the patchiness of reionization (for details see, e.g. ref. 9). Currently the prospects of detecting this signal are being studied by using both single dish and interferometric experiments at low radio frequencies[8] (see also ref. 31).

III. Re-ionization of the universe at high redshift

In the previous section we discussed various probes of the high redshift universe and the epoch of reionization. In this section we take up the issue whether it is possible to explain the observed ionization structure of the universe within the framework of currently-favoured ΛCDM models of formation of structures. In these models, the observed structures in the universe grew from gravitational instability of density fluctuations which originated during inflationary epoch in the very early universe (see e.g. refs 1–3). The gravitational collapse of these structures might either set off star-formation (alternatively some material might end up in black holes which by accreting more matter will radiate with harder spectrum than first star-forming galaxies, see e.g. ref. 32) which will emit UV light and ionize the IGM. The process of re-ionization of the universe is generally quite complicated and not well understood. However it is possible to study it within the framework of simple models which might give important clues about the details of this process. Important ingredients of this problem are: (a) Halo population at high redshift, (b) molecular and atomic cooling in Haloes, (c) initial mass function of stars and star formation rate, (d) Escape fraction of UV photons from Haloes, (e) clumpiness of the IGM. Of these the most uncertain are (c) and (d) and have to be modelled using simple parameterized model.

Halo population

The number density of dark matter haloes per unit mass M at any redshift can be obtained from the Press–Schechter method (see e.g. refs 1–3). It is given by:

$$\frac{{\rm d}n}{{\rm d}M} = \sqrt{\frac{2}{p}}\,\frac{r_m}{M}\,d_c(z)\left|\frac{{\rm d}s}{{\rm d}M}\right|\frac{1}{s^2(M)}\exp[-d_c(z)^2/2s^2(M))]\,. \tag{11}$$

Here $s(M)$ is the mass dispersion filtered at the scale corresponding to mass M in the linear theory; $s(M)$ for length scale corresponding to $8\,{\rm h}^{-1}$ MPc is $\simeq 0.9$ (ref. 14); $d_c(z) \simeq 1.7D(0)/D(z)$, with $D(z)$ being the solution of growing mode in linear theory; in sCDM model $D(0)/D(z) = (1 + z)$ (for details see refs 2, 3, 7). The first baryonic structures to collapse would have masses exceeding the Jeans mass of the IGM, whose evolution is shown in Figure 1. From Figure 1 it is seen that the first structures might have masses $\simeq 10^4\,M_\odot$ (for nonlinear extension of the concept of Jeans mass see e.g. ref. 4; the modified mass has values similar to the one obtained using linear theory). The collapsed fraction of all structures can be calculated from eq. (11) and is $\simeq 2 \times 10^{-3}$ at $z \simeq 20$. However to form stars, baryons need to cool sufficiently rapidly in the dark matter haloes to collapse to density higher than initial virialized density.

Atomic and molecular cooling

Assuming a spherical, top-hat collapse (see e.g. refs 2, 3) the density of the collapsed structure is $\simeq 170$ times the background density at that redshift. If a virial equilibrium is reached inside the halo, the baryon temperature is raised to the virial temperature given by (e.g. refs 2, 3):

$$T_{\rm vir} \simeq 800{\rm K}\left(\frac{M}{10^6\,M_\odot}\right)^{2/3}\left(\frac{1+z}{20}\right)\left(\frac{\Omega_m}{0.3}\right)^{1/3}\left(\frac{h}{0.7}\right)^{2/3}\left(\frac{m}{1.22}\right). \tag{12}$$

Here molecular weight $m = 1.22$ for a fully neutral halo (haloes with masses $\lesssim 10^8\,M_\odot$ at $z \simeq 20$); $m = 0.57$ for a fully ionized halo. An important criterion is whether baryons can cool rapidly enough so that they collapse to higher densities, fragment and form stars. In the primordial gas, the cooling in haloes with virial temperature $\gtrsim 10^4$ is dominated by atomic hydrogen and singly-ionized helium. But the smaller haloes can only cool further by molecular hydrogen. A small fraction $\simeq 10^{-6}$ of molecular hydrogen is formed at $z \simeq 1000$ in the IGM following recombination of the universe (see e.g. ref. 1). Such a small fraction does not suffice to cause rapid enough cooling. However the collapse of halo can cause the formation of molecular hydrogen, resulting in a molecular fraction $\simeq 5 \times 10^{-4}$ by the time the halo virializes (see e.g. ref. 23 and references therein). In Figure 7, we show the cooling time (defined as $t_{\rm cool} = kT_K/\dot{E}$, where $\dot{E}$ is the rate at which the halo loses energy from various processes) for haloes of different masses at $z = 20$. The criterion for runaway cooling resulting in fragmentation and star formation is that cooling time be less than the dynamical time $(t_d = 1/\sqrt{Gr})$ of the halo. From Figure 7 it is seen than haloes of masses $\simeq 10^6\,M_\odot$ can collapse to form stars at $z \simeq 20$. This result is only approximate as the haloes will not have constant density. For instance, the more dense central parts of the halo can form both more molecular hydrogen and cool more rapidly. Bromm *et al.*[34] simulated the collapse of haloes of masses $\simeq 10^6\,M_\odot$ with substructure. They concluded that the halo fragments into many clumps with typical masses 10^2–$10^3\,M_\odot$; their analysis suggests that this mass scale is a result of molecu-

lar hydrogen chemistry and therefore should correspond to the masses of the first stars. Abel *et al.*[35] reached similar conclusions from their simulations.

From information about the halo population and cooling arguments it is possible to speculate on the ionization history of the universe from photo-ionization. The main uncertainty is the hydrogen–ionizing luminosity (and its evolution) of each halo, which has to be parameterized. Assuming that halo of mass M emits isotropically the hydrogen-ionizing luminosity $\dot{N}_g$ (in photons sec^{-1}), the radius of ionizing sphere around the source will satisfy the equation (Stromgren Sphere, see e.g. refs 36, 37):

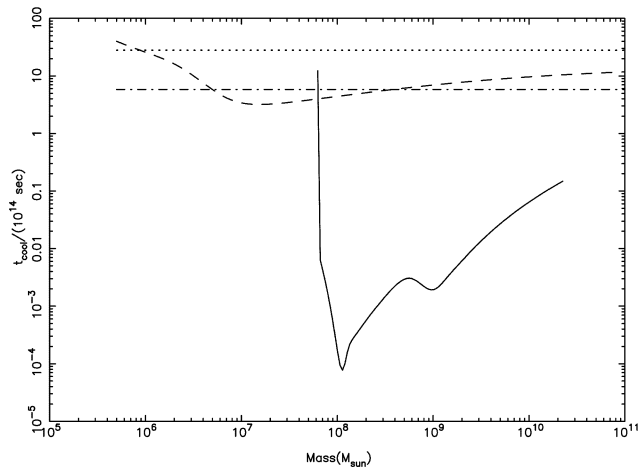

Figure 7. Cooling time is shown against the mass of the virialized halos at $z = 20$. The solid line is from atomic cooling of the primordial gas containing only hydrogen and helium. The cooling processes include: line cooling from hydrogen and helium, free–free emission, and recombination. The dashed curve is for cooling from molecular hydrogen for a molecular fraction 5×10^{-4}. The dotted line is the inverse of local expansion rate $H^{-1}(z)$ and the dot-dashed line is the dynamical time $1/\sqrt{Gr_m}$.

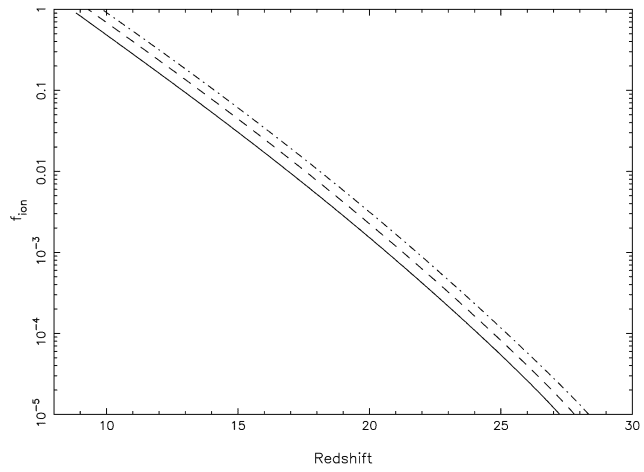

Figure 8. Ionization history is shown for several values of $N_g(0)$ and C (see text for details). Solid, dashed, and dot-dashed curves are: $\{N_g(0) = 10^{50}$ s^{-1}, $C = 1$; $N_g(0) = 2 \times 10^{50}$ s^{-1}, $C = 2$; $N_g(0) = 5 \times 10^{50}$ s^{-1}, $C = 5\}$, respectively.

$$\frac{\mathrm{d}R}{\mathrm{d}t} - HR = \frac{(\dot{N}_g - 4p/3R^3 a_B C n_b^2 x_{\mathrm{HI}})}{(\dot{N}_g + 4p R^2 x_{\mathrm{HI}} n_b)}. \tag{13}$$

Here $C \equiv \langle n_b^2 \rangle / \langle n_b \rangle^2$ is the clumping factor of the IGM. Using eq. (11), the fraction of the universe that is ionized as a given redshift is (see e.g. refs 30, 38):

$$f_{\mathrm{ion}}(z) = \frac{4p}{3} \int_0^z \mathrm{d}z' \int \mathrm{d}M \, \frac{\mathrm{d}n}{\mathrm{d}M}(M, z') R^3(M, z, z'). \tag{14}$$

Further assuming that the luminosity of the source is $\propto M$, the ionized fraction can be calculated in terms of the evolution of the photon luminosity of a single halo of some fiducial mass and the evolution of the clumping factor. For simplicity we assume the clumping factor to have a constant value between one and five, and take the luminosity evolution of a halo to have the form of a typical star-burst galaxy: $\dot{N}_g(t) = \dot{N}_g(0)\exp(-t/10^7$ years). In Figure 8 we show several ionization histories for different values of $N_g(0)$ for $M = 5 \times 10^7 M_\odot$ and clumping factor C. In Figure 8 we only plot ionization fraction up to $z \simeq 9$ at which $f_{\mathrm{ion}} \simeq 1$; further evolution keeps the universe fully ionized up to the present. These ionization histories are consistent with WMAP observations. (If the universe becomes fully ionized at $z \simeq 8$ it will remain ionized up to the present even in the absence of ionizing sources as the typical recombination time $1/(a_B C n_b)$ already exceeds the age of the universe at this epoch for $C \lesssim 2$). It appears that $N_g(0) \simeq 10^{50}$ is required to ionize the universe early enough to satisfy WMAP observations. This is just three orders of magnitude below the photon luminosity of a typical star-burst galaxy (see e.g. ref. 39). Alternatively one can infer that the efficiency of the first star formation was very high (see e.g. refs 38, 40). There are other uncertainties like feedback from supernova, and photo-dissociation of molecular hydrogen which indicate this estimate is a lower limit (see e.g. refs 4, 38, 41, 42). Alternatively it is possible that the collapsed fraction of the universe far exceeded the value given by ΛCDM models and was caused by some other physical process like tangled magnetic fields[43].

Is it possible to simultaneously satisfy the WMAP and the GP observations? It appears difficult to achieve it unless the efficiency of star formation decreased rapidly from $z \simeq 20$ to $z \simeq 6$ (for other possible scenarios and details see e.g. refs 38, 40 and references therein).

IV. Discussion

The ionization history of the universe at high redshift as inferred by WMAP and GP observations is quite complicated: the universe reionized at $z \simeq 12$–15; the neutral fraction increased to better than one part in a thousand for $5.5 < z < 6.5$; and for $z \lesssim 5$, the neutral fraction dropped to one part in a million. Theoretical analyses based on models of

structure formation are not sophisticated enough to understand this ionization history. Future observations however might throw light on these observations. First, future telescopes might be able to detect the sources of reionization[38] up to $z \simeq 20$. Future CMBR experiment Planck also is sensitive enough to disentangle the effects of complicated ionization histories[44].

1. Peebles, P. J. E., *Principles of Physical Cosmology*, Princeton University Press, 1993.
2. Padmanabhan, T., *Theoretical Astrophysics, Volume III: Galaxies and Cosmology*, Cambridge University Press, Cambridge, 2002.
3. Padmanabhan, T., *Structure Formation in the Universe*, Cambridge University Press, Cambridge, 1993.
4. Barkana, R. and Loeb, A., *Phys. Rep*., 2001, **349**, 125.
5. Hu, W. and Dodelson, S., *ARA&A*, 2002, **40**, 171.
6. Bond, R., Schaeffer, R., Silk, J., Spiro, M. and Zinn-Justin, J., *ASP*, 1996.
7. Peebles, P. J. E., *The Large Scale Structure of the Universe*, Princeton University Press, 1980.
8. Shaver, P. A., Windhorst, R. A., Madau, P. and de Bruyn, A. G., *A&A*, 1999, **345**, 380.
9. Tozzi, P., Madau, P., Meiksin, A. and Rees, M. J., *ApJ*, 2000, **528**, 597.
10. Kogut, A. *et al.*, *ApJS*, 2003, **148**, 161.
11. Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L. and Rix, H., *AJ*, 2002, **123**, 1247.
12. Djorgovski, S. G., Castro, S., Stern, D. and Mahabal, A. A., *ApJL*, 2001, **560**, L5.
13. Becker, R. H. *et al*., *AJ*, 2001, **122**, 2850.
14. Spergel, D. N. *et al*., *ApJS*, 2003, **148**, 175.
15. Perlmutter, S. *et al*., *ApJ*, 1999, **517**, 565.
16. Riess, A. G. *et al.*, *AJ*, 1998, **116**, 1009.
17. Tytler, D., O'Meara, J. M., Suzuki, N. and Lubin, D., *Phys. Rep*., 2000, **333**, 409.
18. Freedman, W. L. *et al*., *ApJ*, 2001, **553**, 47.
19. White, R. L., Becker, R. H., Fan, X. and Strauss, M. A., *ApJ*, 2003, **126**, 1.
20. Dodelson, S. and Jubas, J. M., *ApJ*, 1995, **439**, 503.
21. Hu, W., Scott, D. and Silk, J., *Phys. Rev*., 1994, **D49**, 648.
22. Vishniac, E. T., *ApJ*, 1987, **322**, 597.
23. Seljak, U. and Zaldarriaga, M., *ApJ*, 1996, **469**, 437.
24. Hu, W. and Sugiyama, N., *ApJ*, 1995, **444**, 489.
25. Zaldarriaga, M. and Harari, D. D., *Phys. Rev*., 1995, **D52**, 3276.
26. Zaldarriaga, M., *Phys. Rev*., 1997, **D55**, 1822.
27. Field, G., *ApJ*, 1959, **129**, 551.
28. Field, G., *Proc. IRE*, 1958, **46**, 240.
29. Madau, P., Meiksin, A. and Rees, M. J., *ApJ*, 1997, **475**, 429.
30. Sethi, S. K., 2004, in preparation.
31. http://www.lofar.org/
32. Ricotti, M. and Ostriker, J. P., *MNRAS*, 2004, tmp, 130.
33. Tegmark, M., Silk, J., Rees, M. J., Blanchard, A., Abel, T. and Palla, F., *ApJ*, 1997, **474**, 1.
34. Bromm, V., Coppi, P. S. and Larson, R. B., *ApJL*, 1999, **527**, 5.
35. Abel, T., Bryan, G. and Norman, M., *ApJ*, 2000, **540**, 39.
36. Shu, F. H., *Physics of Astrophysics*, *Gas Dynamics*, University Science Books, vol. II, 1992.
37. Shapiro, P. and Giroux, M. L., *ApJ*, 1987, **321**, L107.
38. Haiman, Z. and Holder, G., *ApJ*, 2003, **595**, 1.
39. Leitherer, *et al.*, *ApJS*, 1999, **123**, 3.
40. Chiu, W. A., Fan, X. and Ostriker, J. P., *ApJ*, 2003, **599**, 759.
41. Haiman, Z., Rees, M. and Loeb, A., *ApJ*, 1997, **476**, 458.
42. Dekel, A. and Silk, J., *ApJ*, 1986, **303**, 39.
43. Sethi, S. K. and Subramanian, K., *MNRAS*, 2005, **356**, 778.
44. Kaplinghat, M., Chu, M., Haiman, Z., Holder, G. P., Knox, L. and Skordis, C., *ApJ*, 2003, **583**, 24.