The Epoch of Reionization

Saleem Zaroubi

Abstract The Universe’s dark ages end with the formation of the first generation of galaxies. These objects start emitting ultraviolet radiation that carves out ionized regions around them. After a sufficient number of ionizing sources have formed, the ionized fraction of the gas in the Universe rapidly increases until hydrogen becomes fully ionized. This period, during which the cosmic gas went from neutral to ionized, is known as the Universe’s Epoch of Reionization. The Epoch of Reionization is related to many fundamental questions in cosmology, such as properties of the first galaxies, physics of (mini-)quasars, formation of very metal-poor stars and a slew of other important research topics in astrophysics. Hence uncovering it will have far reaching implications on the study of structure formation in the early Universe. This chapter reviews the current observational evidence for the occurrence of this epoch, its key theoretical aspects and main characteristics, and finally the various observational probes that promise to uncover it. A special emphasis is put on the redshifted 21 cm probe, the various experiments that are currently being either built or designed, and what we can learn from them about the Epoch of Reionization.

1 Introduction

The formation of the first galaxies marks a major transition in the evolution of structure in the Universe. These same galaxies with their zero metallicity Population III stars, second generation Population II stars, and black hole driven sources (e.g., mini-quasars, x-ray binaries, etc.) transformed the intergalactic medium from neu-
tral to ionized. This process, known as the Epoch of Reionization (EoR), is the central topic discussed in this chapter.

As mentioned in chapter # 1 [by A. Loeb in this book], about 400,000 years after the Big Bang, the Universe’s density decreased enough so that the temperature fell below 3000 K, allowing ions and electrons to (re)combine into neutral hydrogen and helium – the fraction of heavier elements was negligible. Immediately afterwards, photons decoupled from baryons and the Universe became transparent, leaving a relic signature known as the cosmic microwave background (CMB) radiation. This event ushered the Universe into a period of darkness, known as the dark ages.

The dark ages ended about 400 million years later, when the first galaxies formed and start emitting ionizing radiation. Initially during the EoR, the intergalactic medium (IGM) is neutral except in regions surrounding the first objects. However, as this reionization progresses, an evolving patchwork of neutral (H I) and ionized hydrogen (H II) regions unfolds. After a sufficient number of UV-radiation emitting objects formed, the temperature and the ionized fraction of the gas in the Universe increase rapidly until eventually the ionized regions permeate to fill the whole Universe [9, 115, 30, 44, 41, 67, 135].

Fig. 1 This figure shows a slice through redshift of the 21 cm radiation in which the reionization process progresses through the volume of a cosmological simulation with radiative transfer [206].

The current constraints strongly suggest that the EoR roughly occurs within the redshift range of $z \sim [6 - 15]$. Figure [1] shows a space-redshift slice of a simulation of the progression of reionization with time and how it appears in 21 cm brightness temperature, which is proportional to the density of neutral hydrogen (see section 4). At high redshifts most of the gas is neutral, hence, the signal is mostly sensitive to cosmological density fluctuations, whereas at lower redshifts ionization bubbles start to appear until they fill the whole Universe [9].

The EoR is a watershed epoch in the history of the Universe. Prior to it, the formation and evolution of structure was dominated by dark matter alone, while baryonic matter played a marginal role. The EoR marks the transition to an era in which the role of cosmic gas in the formation and evolution of structure became prominent and, on small scales, even dominant.
The details of the reionization scenario I have laid out are yet to be clarified. For example, it is not known what controls the formation of the first objects and how much ionizing radiation they produce, or how the ionization bubbles expand into the intergalactic medium and what they ionize first, high-density or low-density regions? The answer to these questions and many others that arise in the context of studying the EoR needs knowledge of fundamental issues in cosmology, galaxy formation, quasars and the physics of very metal poor stars; all including foremost research in topics in modern astrophysics. Substantial theoretical and observational efforts are currently dedicated to understanding the physical processes that trigger this epoch and govern its evolution, and ramifications on subsequent structure formation (c.f., [9, 30, 44, 41, 67]). However, despite the pivotal role played by the EoR in cosmic history, observational support for the proposed scenarios is very scarce, and when available, is indirect and model dependent.

In principle, there are many different ways to observationally probe the EoR. In this contribution, I mainly focus on the redshifted 21cm emission line from neutral hydrogen at high redshifts. This is one of the most promising techniques for studying the dark ages and the EoR. To date, there are a number of telescopes dedicated to measure this faint radiation. In the short term, these consist of: The Low Frequency Array (LOFAR), the Murchison Widefield Array (MWA), Precision Array to Probe Epoch of Reionization (PAPER) and Giant Metrewave Radio Telescope (GMRT), while, on a somewhat longer time scales the Square Kilometer Array (SKA). One of the most challenging tasks in studying the EoR is to extract and identify the cosmological signal from the data and interpret it correctly. This is because the detectable signal in the frequency range relevant to the EoR is composed of a number of components – the cosmological EoR signal, extragalactic and Galactic foreground, ionospheric distortions, instrumental response and noise – each with its own physical origin and statistical properties.

Figure 2 shows a sketch of the likely evolution of reionization from the neutral hydrogen point of view. The figure emphasizes the other non-cosmological effects that are seen with the 21 cm experiments, e.g., foreground, ionosphere and instrumental effects. The radio antennas seen at the bottom are LOFAR’s Low Band Antennas.

In this chapter I discuss various observational and theoretical aspects of the Epoch of Reionization. In section 2 the current observational scene is reviewed, specifically focusing on the CMB data and the Lyman α forest spectra. In sections 3 and 4, we discuss, respectively, the physics of the reionization process and the 21 cm line transition and how it could be used to probe reionization. The redshifted 21 cm experiments their potentials and the challenges are discussed in section 5. Extraction and quantification of the information stored in the redshifted 21 cm data using various statistics is discussed in section 6. This chapter concludes with a brief summary (section 7).
Fig. 2 This figure shows a sketch of the likely development of the EoR. About 500,000 years after the Big Bang ($z \sim 1000$) hydrogen recombined and remained neutral for a few hundred million years during the dark ages. At a redshift, $z \sim 15$, the first stars, galaxies and quasars began to form, heating and reionizing the hydrogen gas. The neutral IGM can be observed with LOFAR up to $z \approx 11.5$ through its redshifted 21cm spin-flip transition. However, many atmospheric, galactic and extra-galactic emission contaminate the 21 cm signal.

2 Observational Evidence for Reionization

To date, the majority of observations related to the EoR provide weak and model dependent constraints on reionization. However, there are currently a number of observations which could impose strong constraints on reionization models, as dis-
The state of the intergalactic medium (IGM) can be studied through the analysis of the Lyman-α forest. This is an absorption phenomenon seen in the spectra of background quasi-stellar objects (QSOs). The history of this field goes back to 1965 when a number of authors [74, 179] predicted that an expanding Universe, homogeneously filled with gas, will produce an absorption trough due to neutral hydrogen, known as the Gunn-Peterson trough, in the spectra of distant QSOs bluewards of the Lyman-α emission line of the quasar. That is, the quasar flux will be absorbed at the UV resonance line frequency of 1215.67 Å. Gunn & Peterson [74] found such a spectral region of reduced flux, and used this measurement to put upper limits on the amount of intergalactic neutral hydrogen. The large cross-section for the Lyman α absorption makes this technique very powerful for studying gas in the intergalactic medium.

In the last 15 years two major advances occurred. The first was the development of high-resolution echelle spectrographs on large telescopes (e.g., HIRES on the Keck and UVES on the Very Large Telescope) that provided data of unprecedented quality. The second was the emergence of a theoretical paradigm within the context of cold dark matter (CDM) cosmology that accounts for all the features seen in these systems (e.g. [16, 39, 81, 118, 131, 201, 232, 233]). According to this paradigm, the absorption is produced by volume filling photoionized gas that contains most of the baryons at redshifts at $z \sim 3 − 6$ and resides in mildly non-linear overdensities.

Figure 3 shows a typical example of the Lyman α forest seen in the spectrum of the $z = 3.12$ quasar Q0420-388. An interesting feature of such spectra is the density of weak absorbing lines which increase with redshift due to the expansion of the Universe. In fact, at redshifts above 4, the density of the absorption features become so high that it is hard to define them as separate absorption features. Instead, one sees only the flux in between the absorption minima which appears as if they are emission rather than absorption lines.

The Lyman α forest has turned out to be a treasure trove for studying the intergalactic medium and its properties in both low and high density regions. In particular, it is very sensitive to the neutral hydrogen column density and hence, to the neutral fraction as a function of redshift along the line of sight. In the following, we demonstrate how one could constrain the neutral fraction of hydrogen from the forest and what the values obtained from the data are. For a review on the Lyman α forest the reader is referred to [162].

We need to calculate the optical depth for absorption of Lyman α photons. A photon emitted by a distant quasar with an energy higher than 10.196 eV is continuously redshifted as it travels through the intergalactic medium until it reaches the observer. At some intermediate point the photon is redshifted to around 1216 Å in
The rest-frame of the intervening medium, which may contain neutral hydrogen. It can then excite the Lyman $\alpha$ transition and be absorbed. Let us consider a particular line of sight from the observer to the quasar. The optical depth $\tau_{\alpha}$ of a photon is related to the probability of the photon’s transmission $e^{-\tau_{\alpha}}$. At a given observed frequency, $v_0$, the Lyman $\alpha$ optical depth is given by

$$\tau_{\alpha}(v_0) = \int_0^{Q} n_{HI}\sigma_{\alpha}dl/(1+z),$$

(1)

where $l$ is the comoving radial coordinate of some intermediate point along the line of sight, $z$ is the redshift and $n_{HI}$ is the proper number density of neutral hydrogen at
The Epoch of Reionization

that point. The limits of the integration, \(O\) and \(Q\), are the comoving distance between the observer and the quasar, respectively. The Lyman \(\alpha\) absorption cross section is denoted by \(\sigma_\alpha\). It is a function of the frequency of the photon, \(\nu\), with respect to the rest-frame of the intervening \(H i\) at position \(l\). The cross section is peaked when \(\nu\) is equal to the Lyman \(\alpha\) frequency \(\nu_\alpha\). The frequency \(\nu\) is related to the observed frequency \(\nu_0\) by \(\nu = \nu_0(1+z)\), where \(1+z\) is the redshift factor due to the uniform Hubble expansion alone at the same position. Note that for the sake of simplicity here we ignore peculiar velocity effects.

Using \(dl = cdt/a\), where \(a\) is the Hubble scale factor and \(t\) is the proper time and Friedmann equation for a flat Universe with cosmological constant, we have,

\[
\tau_\alpha = \int \sigma_\alpha(\nu)n_{HI} \frac{cH_0^{-1}dz}{(1+z)^{\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}}}. \tag{2}
\]

This optical depth should also depend on the Lyman \(\alpha\) line profile function but here we assume that it is basically a \(\delta\)-function centered at the frequency \(\nu\). Considering \(n_{HI} = n_Hx_{HI}\), where \(x_{HI}\) is the neutral fraction of hydrogen, and integrating over this equation, one obtains the following result:

\[
\frac{n_{HI}}{n_H} = x_{HI} \approx 10^{-4}\Omega_m^{1/2}h(1+z)^2 \tau_\alpha. \tag{3}
\]

Since the Lyman \(\alpha\) features mostly show mild absorption probability (\(\tau_\alpha \lesssim 1\)) this equation clearly implies that at the mean density of the Universe at \(\tau_\alpha\) of about one the ionized fraction is on the order of \(10^{-4}\). Therefore, the fact that we observe the Lyman \(\alpha\) forest at all means that the Universe is highly ionized at least until \(z \approx 6\). This is the most reliable and robust evidence that the Universe has in fact reionized.

Another important evidence relevant for reionization comes from high resolution spectroscopy of high redshift Sloan Digital Sky Survey (SDSS) quasars [59, 60]. The SDSS has discovered about 19 QSOs with redshifts around 6 that are powered by black holes with masses on the order of \(10^9 M_\odot\). In a follow up observations with 10 meter class telescopes Fan et al. [59, 60] were able to obtain high resolution spectra of these objects.

Fig. 4 shows the spectra of these high redshift quasars [59, 60]. Notice the complete absence of structure that some of these spectra exhibit bluewards of the quasar Lyman \(\alpha\) restframe emission, especially those with redshift \(z \gtrsim 6\). This is normally attributed to an increase in \(\tau_\alpha\) as a result of the decrease in the ionized fraction of the Universe. Notice also, that although the trend with redshift is clear, it is by no means monotonic. For example, quasar J1411+3533 at \(z = 5.93\) shows an “emptier” trough relative to quasars J0818+1722 at \(z = 6\). Such trend might be indicating a more patchy ionization of the IGM at such redshifts.

Figure 5 shows the effective Lyman \(\alpha\) or Gunn-Peterson optical depth, \(\tau_{\text{eff}}\), as a function of redshift as estimated from the joint optical depths of Lyman \(\alpha\), \(\beta\) and \(\gamma\). From this plot it is clear that the increase in the optical depth as a function of redshift is much larger than expected (shown in the dashed line) from passive redshift evolution of the density of the Universe.
Fig. 4 Spectra for high redshift SDSS quasars. The Gunn-Peterson trough bluewards of the QSO Lyman α emission that is clearly apparent in the highest redshift ones indicates that the Universe has become somewhat more neutral at these redshifts. A similar behavior is also seen bluewards of the QSO Lyman β region of the same spectra. The actual amount of increase in neutral hydrogen implied by these spectra is not clear [60].
The interpretation of the increase in the optical depth at $z \gtrsim 6.3$ has been the subject of some debate. All authors agree that this is a sign of an increase in the Universe’s neutral fraction at high redshifts, marking the tail end of the reionization process. The controversy is centered on the question of by how much the neutral fraction increases. Some authors \cite{217,218,130} have argued that the size of the so-called Near Zone ionized by the quasar itself and set redwards of the Gunn-Peterson trough indicates that the neutral fraction around the SDSS high redshift quasars is $\approx 10\%$. More recently, it has been suggested that the variations seen across various SDSS quasars indicate that the ionization state of the IGM at these redshifts changes significantly across different sightlines \cite{128}. However, given the intense radiation field around these quasars, it is not possible to put general constraints on the neutral fraction of the IGM from quasars at redshift below 6.5 (see e.g., \cite{18,216,123,124}). Moreover, recently and with the discovery of the redshift $z = 7.1$ QSO ULAS J1120+0641 \cite{137} by the UKIDSS survey \cite{108} it has been argued that this quasar’s Near Zone gives a clear evidence for an increase in the neutral fraction of hydrogen in the IGM at $z = 7.1$ \cite{137,20}. Note however that his conclusion relies on one quasar and might change as more of such quasars at $z \gtrsim 7$ are discovered.

There are more things that we can learn about reionization from the Lyman $\alpha$ forest that we will discuss later. But to summarize, the main conclusion from the Lyman $\alpha$ optical depth measurements is that the Universe is highly ionized at redshifts below 6 (as seen in Figure 4), while at about $z=6.3$, its neutral fraction increases, forming the tail end of the reionization process (see Figure 5).

### 2.2 The Thomson Scattering Optical depth for the Cosmic Microwave Background (CMB) Radiation

This is a very evolved topic, discussed and reviewed by many authors (e.g., \cite{157,195,21,117,86,4}). Here, I give a general review of the constraints provided by the CMB on reionization. The CMB provides important information relevant to the history of reionization. It is known that the Universe has indeed recombined and became largely neutral at $z \approx 1100$. If recombination had been absent or substantially incomplete, the resulting high density of free electrons would imply that photons could not escape Thomson scattering until the density of the Universe dropped much further. This scattering would inevitably destroy the correlations at subhorizon angular scales seen in the CMB data (see e.g., \cite{85,190}).

In order to calculate the effect of reionization on CMB photons, a function is often defined called the visibility function\footnote{Notice that this is a different “visibility” than the one used in radio interferometry which we discuss in section 5.}

$$g(\eta) = -\tau e^{-\tau(\eta)},$$

(4)
Evolution of the Lyman $\alpha$, $\beta$ and $\gamma$ optical depth from the high redshift Sloan quasars. The Lyman $\beta$ and Lyman $\gamma$ restframe wavelengths are 1026 Å and 972.5 Å, respectively. The Lyman $\beta$ measurements are converted to Lyman $\alpha$ Gunn-Peterson optical depth using a conversion factor that reflects the difference in the cross section between the two transitions, which is a factor of 5.27 lower in the case of Lyman $\beta$ (see [112, 38]). The values in the two highest redshift bins are lower limits, since they both contain complete Gunn-Peterson troughs. The dashed line shows a redshift evolution of $\tau_\alpha \approx (1 + z)^{4.3}$. At $z > 5.5$, the best-fit evolution has $\tau_\alpha \approx (1 + z)^{10.9}$, indicating an accelerated evolution. The large filled symbols with error bars are the average and standard deviation of the optical depth at each redshift. The sample variance also increases rapidly with redshift. Figure taken from [60].

where $\eta (\equiv \int dt / a)$ is the conformal time, $a$ is the scale factor of the Universe and $\dot{\tau}$ is the derivative of the optical depth with respect to $\eta$. The optical depth for Thomson scattering is given by $\tau(\eta) = - n_e(\eta) \int_0^\eta \frac{d\eta'}{a(\eta')} \sigma_T$, where $n_e$ is the electron density and $\sigma_T$ is the Thomson cross section. The visibility function gives the probability density that a photon had scattered out of the line of sight between $\eta$ and $\eta + d\eta$. The influence of reionization on the CMB temperature fluctuations is obtained by integrating Equation 4 along each sightline to estimate the temperature fluctuation suppression due to the EoR. The suppression probability turns out to be roughly proportional to $1 - e^{-\tau}$ [224]. Since the amount of suppression in the measured power spectrum is small, the optical depth for Thomson scattering must be small too [152]. The left hand panel in Figure 6 shows the influence of increasing the value of $\tau$, the Thomson optical depth, on the CMB temperature fluctuation power spectrum. The right hand panel shows the reionization history of the Universe assumed in the left panel. Since in this case a
The Epoch of Reionization

Fig. 6  Left hand panel (a): The influence of reionization on the CMB temperature angular power spectrum. Reionization damps anisotropy power as $e^{-2\tau}$ under the horizon (diffusion length) at last scattering. The models here are fully ionized $x_e = 1.0$ out to a reionization redshift $z_i$. Notice that with high optical depth, fluctuations at intermediate scales are regenerated as the fully ionized (long-dashed) model shows. This figure is taken from Wayne Hu’s PhD thesis [85]. Right panel (b) shows the assumed reionization history used. It is obvious that since we are considering a uniform and sudden reionization model, a change in the reionization redshift, $z_i$, will translate uniquely to an optical depth for Thomson scattering.

sudden global reionization is assumed, there is one to one correspondence between the optical depth for Thomson scattering and the redshift of reionization.

Fig. 7  Left hand panel: A sketch that shows why the CMB polarization is sensitive to the quadrupole momentum of temperature fluctuations. Right hand panel: Thomson scattering of radiation with quadrupole anisotropy generates linear polarization. The blue and red lines represent cold and hot radiation.

Further information can be obtained from observations of CMB via the polarization power spectrum. The polarization of the CMB emerges naturally from the Cold Dark Matter paradigm which stipulates that small fluctuations in the early universe grow, through gravitational instability, into the large scale structure we see today ([21][85][97][227]). Since, the temperature anisotropies observed in the CMB are the result of primordial fluctuations, they would naturally polarize the CMB.
anisotropies. The degree of linear polarization of the CMB photons at any scale reflects the quadrupole anisotropy in the plasma when they last scattered at that same scales. From this argument it is clear that the amount of polarization at scales larger than the horizon scale at the last scattering surface should fall down since there is no more coherent quadrupole contribution due to the lack of causality. This is shown in the sketch presented in the left hand panel in Figure 7. The largest scale at which a primordial quadrupole exists is the scale of the horizon at recombination, which roughly corresponds to 1°. Therefore, any polarization signature on scales larger than the horizon scale provides a clear evidence for Thomson scattering at later stages where the horizon scale is equivalent to the scale on which polarization has been detected.

Furthermore, the polarized fraction of the temperature anisotropy must be small, normally one order of magnitude smaller than the anisotropy in the temperature. This is simply because these photons must have passed through an optically thin plasma, otherwise they would not have reached us but they would have scattered and destroyed the sub-horizon (i.e., below 1°) correlation in the CMB, contrary to what we observe (see e.g., [190]).

The dependence Thomson scattering differential cross section on polarization is expressed as

\[ \frac{d\sigma_T}{d\Omega} = \frac{e^4}{m_e^2 c^4} |\mathbf{e} \cdot \mathbf{e}'|^2 \]  

(5)

where \( e \) and \( m_e \) are the electron charge and mass and \( \mathbf{e} \cdot \mathbf{e}' \) is the angle between the incident and scattered photons. The right hand panel of Figure 7 shows how the Thomson scattering produces polarization of the CMB photons. If the CMB photons scatter later due to reionization and the incident radiation has a quadrupole moment, then it will be scattered in a polarized manner on the scale roughly equivalent to the horizon scale at the redshift of scattering. That is why the scale at which the large scale polarization is detected gives information about the reionization redshift.

The polarization field of the CMB photons is usually described in terms of the so called “electric” (E) and “magnetic” (B) components which can be derived from a scalar or vector field. The harmonics of an E-mode have \((-1)^\ell\) parity on the sphere, whereas those of B-mode have \((-1)^{\ell+1}\) parity. Under parity transformation, i.e., \( \hat{n} \rightarrow -\hat{n} \), the E-mode thus remains unchanged for even \( \ell \), whereas the B-mode changes sign and vice versa. Fig. 8 illustrates such (a)symmetry under parity transformation for the simple case of \( \ell = 2, m = 0 \) [86].

Various physical processes lead to different effects on the CMB polarization. Most of these effects are expected to produce E mode polarization patterns on the CMB. However, gravitational waves in the primordial signal and gravitational lensing of the CMB on its way to us produce a B mode polarization patterns. A large scale E mode polarization signal could only be caused by the process of reionization. The main reason for this is that large scale polarization could not be caused by causal effects on the last scattering surface which has a 1° scale whereas reionization, which occurs much later, has no such restriction. Figure 7 shows the measured and predicted CMB angular power and cross-power spectra from the WMAP 3rd
year data. The existence of large scale correlation in the E-mode is a strong indication that the Universe became ionized around redshift $z \approx 10$. The argument in essence is mostly geometric, namely it has to do with the scale of the E-mode power spectrum as well as the line of sight distance to the onset of the reionization front along a given direction. Some authors have also argued that one can have somewhat more detailed constraints on reionization from the exact shape of the CMB E-mode polarization large scale bump \cite{84,110,138}. Unfortunately however, the large cosmic variance at large scales limits the amount of possible information one can extract. Still, the Planck surveyor is expected to be able to retrieve some of the large scale bump shape.

From Figure 8 one can also deduce the optical depth for Thomson scattering, $\tau$, caused by the scattering of the CMB photons off free electrons released by reionization to be $0.087 \pm 0.017$ \cite{57}. This could be turned into a constraint on the global reionization history through the integral,

$$\tau = \int_{z_{\text{dec}}}^{\infty} \sigma_T n_e \frac{cH_0^{-1}dz}{(1+z)\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}}. \quad (6)$$

Here $z_{\text{dec}}$ is the decoupling redshift, $\sigma_T$ is the Thomson cross section, $\mu$ is the mean molecular weight and $n_e$ is the electron density. This formula works for the optical depth along each sight line but also for the mean electron density, i.e., mean reionization history, of the Universe.

An important point to notice here is that, in order to turn $\tau$ into a measurement of the reionization redshift, one needs a model for $n_e$ as a function of redshift. Hence, one has to be careful when using the reionization redshift given by CMB papers as in most cases a gradual reionization is assumed. Sudden reionization gives a
Fig. 9 The temperature and E-mode polarization power and cross-power spectra as measure by the WMAP satellite [152]. Plots of signal for TT (black), TE (red), and EE (green) for the best-fit model. The dashed line for TE indicates areas of anticorrelation. For more details about this Figure we refer the reader to the Page et al. [152]. Notice the excess power on large scales caused by reionization seen in the TE and EE power spectra.

one to one correspondence between the measured optical depth and the reionization redshift, e.g., the WMAP measurement optical depth implied $z_i = 11.0 \pm 1.4$.

However, sudden reionization is very unlikely and most models predict a more gradual evolution of the electron density as a function of redshift. Furthermore, in such scenarios the redshift of reionization is not clearly defined, therefore authors refer instead to the redshift at which half of the IGM volume is ionized, $z_{\text{Hi} \text{I}}=0.5$. Obviously, in the case of sudden reionization the two redshifts coincide, $z_i = z_{\text{Hi} \text{I}}=0.5$.

It is also important to notice that in the case of sudden reionization the WMAP measured Thomson optical depth does not imply that the redshift at which half the IGM is ionized is the same as $z_i$, and in most cases one obtains $z_{\text{Hi} \text{I}}=0.5 < z_i$ [206].

The patchy nature of the reionization process will also leave an imprint at arcminute scales on the CMB sky. Such an imprint will be mostly caused by the reionization bubbles that form during the EoR. However, the strength of the reionization signal at small scales is found to be smaller than that caused by gravitation lensing.
and is very hard to extract unless the experiment has a very high signal-to-noise at such small scales [54].

2.3 The Intergalactic Medium at $z \lesssim 6$

There are a number of other observations that put somewhat less certain constraints on reionization. Those constraints come mostly from detailed analysis of high resolution Lyman $\alpha$ forest data and from the observation of high redshift Lyman break galaxies. Here we present the two "strongest" of those constraints.

2.3.1 IGM Temperature Evolution

![Fig. 10](image-url) The different baryon phases in the $\rho - T$ diagram. Gray contours show a mass-weighted histogram: the baryon mass fraction at a given density and temperature. Each region corresponds to a given phase (diffuse background, hot, or cold gas) [161].

Another constraint on the reionization history comes from studying the thermal history of the IGM. Due to its low density, the intergalactic medium cooling time is long and retains some memory of when and how it was last heated, namely, reion-
ized. Hence, measuring the IGM temperature at a certain redshift (\( \gtrsim 3.5 \)) allows us to reconstruct, under certain assumptions, its thermal history up to the reionization phase where the IGM has been substantially heated. Such a measurement has been carried out by a number of authors using high resolution Lyman \( \alpha \) forest data, especially using the very low column density absorption lines. The width of these absorption features carries information about the temperature of the underlying IGM. This temperature obviously varies with density and with other parameters like the background UV flux. Based on both theoretical arguments \([87]\) and on numerical simulations \([201]\) in the linear and quasilinear regime, the temperature-density relation follows the simple power law,

\[
T = \bar{T} \left( \frac{\rho}{\bar{\rho}} \right)^{\gamma - 1},
\]

where \( \bar{T} \) is the temperature of the IGM at the mean density of the Universe and \( \gamma \) is the adiabatic power law index. Figure \([10]\) shows the so called phase diagram, i.e., the relation between the temperature and density, obtained from a cosmological hydrodynamical simulation \([161]\). The relation between the density and temperature at the low density end of the diagram, marked as diffuse background, follows a power law. The hot phase at intermediate densities where cooling is not efficient, is driven by shock heating. At high densities, cooling becomes very efficient and drives the gas temperature. At high redshifts more than 90\% of the gas is in the diffuse phase.

Given the validity of equation \([7]\) at low densities, it is meaningful to define an IGM temperature as the gas temperature at the mean density, \( \bar{T} \). Such a measurement has been performed by a number of authors at \( z \approx 3 - 4 \) \([111, 178, 203, 225]\) and recently at \( z \approx 6 \) by \([17]\).

The usefulness of this temperature to constrain the reionization history was first realized by \([202, 88]\) who used the measured temperature around redshift 3 to set \( z \approx 9 \) as an upper limit for the reionization process. Bolton et al. \([17]\) have recently confirmed these findings with higher redshift quasars. That is, the measured temperatures of the IGM at redshift \( z \approx 3 \) and \( z \approx 6 \) are too high for the bulk of reionization to have occurred at redshift \( \gtrsim 10 \).

After reionization, the evolution of the IGM mean temperature \( \bar{T} \) is given by

\[
\frac{1}{\bar{T}} \frac{d\bar{T}}{dt} - \frac{1}{\mu} \frac{d\mu}{dt} = -2H + \frac{\mu \Delta_e}{2 k_B \bar{T}},
\]

where \( H \) is the Hubble parameter, \( k_B \) is the Boltzmann constant, \( \mu \) is the mean molecular weight, and \( \Delta_e \) is the effective radiative cooling rate (in units of ergs g\(^{-1}\) s\(^{-1}\)). \( \Delta_e \) is negative (positive) for net cooling (heating) and includes photoelectric heating and cooling via recombination, excitation, inverse Compton scattering, collisional ionization, and bremsstrahlung. Without cooling/heating processes the cooling rate is set by adiabatic cooling, namely, Hubble expansion. This equation enables us to calculate the temperature evolution as a function of redshift. Measure-
ing the IGM temperature at a given redshift will allow us to extrapolate back in
time until we reach a temperature of $6 \times 10^4$ K which is the temperature at which
hydrogen ionizes. Figure 11 demonstrates this procedure [202].

Obviously, the weak point of this argument is the assumption that one knows the
cooling/heating function of the IGM at every redshift up to the time of reionization.
Still, this is a useful argument and certainly any model for the reionization history
would have to explain the temperature we measure at lower redshifts.

---

**Fig. 11** Temperature evolution of the IGM above redshift 3.4. The solid curves indicate the
evolution of the temperature at the mean density for various $\text{H} I$ reionization redshifts $z_{\text{H}}$, as
indicated. The temperature after hydrogen reionization is assumed to be $T_0 = 6 \times 10^4$ K, and
the hydrogen photoionization rate is $\Gamma = 10^{-13} s^{-1}$ ($\Gamma = 10^{-14} s^{-1}$, short-dashed curve). The
$\text{He} II$ photoionization rate is adjusted so that the $\text{He} III$ abundance is $x_{\text{HeIII}} \approx 0.1$ at $z = 3.5$. The
solid curve connecting the filled squares indicates $z_{\text{H}} = 10.2$ and a higher $\text{He} II$ photoionization
rate, $x_{\text{HeIII}}(z = 3.5) = 0.6$. Finally, the long-dashed curves has $z_{\text{H}} = 20$ but a still higher
$\text{He} II$ photoionization rate, $x_{\text{HeIII}}(z = 3.5) = 0.95$. If He is mostly singly ionized at $z \approx 3.5$, then the
rapid decrease in $T_0$ after reionization places an upper limit of $z_{\text{H}} < 9$ on the redshift of hydrogen
reionization. The filled squares with error bars show the measured IGM temperature as a function
of redshift. This figure taken from [202].
2.3.2 Number of Ionizing Photons per Baryon

Another constraint that comes mostly from the Lyman $\alpha$ forest but also from the recently discovered galaxies at $z \gtrsim 7$ is the number of ionizing photons per baryon. Using physically motivated assumptions for the mean free path of ionizing photons, Bolton and Haehnelt ([18]) turned the measurement of the photoionization rate into an estimate of the ionizing emissivity. They showed that the inferred ionizing emissivity in comoving units, is nearly constant over the redshift range $2 - 6$ and corresponds to $1.5 - 3$ photons emitted per hydrogen atom over a time interval corresponding to the age of the Universe at $z = 6$. Completion of reionization at or before $z = 6$ requires therefore, either an emissivity which rises towards higher redshifts or one which remains constant but is dominated by sources with a rather hard spectral index, e.g., mini-quasars.

With the installation of the WFC3 camera aboard the Hubble Space Telescope, searches for high redshift galaxies at $z = 6 - 10$ have improved dramatically. In particular, a number of authors [145, 23, 33, 125] have reported detection of very high redshifts galaxies using the Lyman-break drop-out technique. The most striking result of these studies is the low number of galaxies found beyond redshift $\approx 6$, making it very hard for these galaxies to ionize the Universe. This conclusion depends however on assuming a luminosity function for galaxies at these redshifts, a function that is very poorly known. More surprising is the very steep drop in the number of galaxies at redshift $z \approx 9$ [22] which makes it even harder to explain reionization with such galaxies.

The last two observational findings have led some authors to claim that the reionization is photon starved, i.e., has a low number of ionizing photons per baryon, which results in a very slow and extended reionization process [18, 34]. Figure 12 shows the number density of ionizing photons (left-hand vertical axis) and number of ionizing photons per baryon (right-hand vertical axis) as a function of redshift. The number of ionizing photons per baryon at redshift 6 is of the order of 2. More recent results deduced from Lyman-break galaxies are consistent with this figure and show an even lower ratio of ionizing photons per baryon at higher redshifts.

2.4 Other Observational Probes

In addition to the probes that we discussed so far, there are a large number of other observational probes that could potentially add valuable input to the reionization models. Examples of such probes are cosmic infrared and soft x-ray backgrounds [52], Lyman $\alpha$ emitters [149], high redshift QSOs [137] and GRBs [31], metal abundance at high redshift [172], etc. However, such probes currently provide very limited constraints on the EoR.

In the coming chapters we will focus on the very large effort currently made to measure the diffuse neutral hydrogen in the IGM as a function of redshift up to
The Epoch of Reionization

3 The Reionization process

The inflationary process that occurred very early in the Universe has created the initial tiny fluctuations in matter density field. The high density peaks in these fluctuations field are the seeds around which galaxies form. The formation process is initially driven by gravitational instability alone but later gas physics, cooling, heating, radiation processes and feedback effects play an important role as well \cite{132,156}. The first galaxies form when primordial gas (H\textsc{i} and He\textsc{i}) condenses within dark matter potential wells which leads to radiative cooling driven mostly by the Lyman $\alpha$ line transition \cite{53,75,107,106,105,153} and, probably, by H$_2$ cooling. To date Lyman $\alpha$ emission has been observed in many high redshift galaxies \cite{98,140,149,150}. This gas condenses further to form the first stars and black holes which in turn produce radiation that starts ionizing the Universe. The efficiency with
which these objects produce ionizing radiation is subject to many different physical processes and assumptions (see e.g., [44]). Since this book’s topic is the first galaxies, the reader is referred to the other chapters in this volume for detailed discussion of how the first radiation emitting objects form and how efficient are they in producing ionizing radiation.

An important unknown in these galaxies is the so called escape fraction, namely the fraction of ionizing radiation that escapes the galaxy into the IGM. It is these ionizing photons that are relevant to the Universe’s reionization. Determining the escape fraction of ionizing radiation observationally is very difficult especially at high redshifts where the available information is very limited. Nevertheless, such observations have been carried out by a number of authors [70, 91, 92, 181, 189] where the measured fraction is found to be between 0.1-0.5. Theoretical prediction of the escape fraction is also difficult. Early studies have assumed idealized smooth galaxies [55, 56, 167, 214] but later studies have simulated more realistic galaxies (see e.g., [43]). Each of these studies have considered different set up and different sources but all conclude that the escape fraction of radiation is roughly in the range of 0.1-0.5.

The most accepted picture of how reionization unfolds is simple. The first radiation-emitting objects ionize their immediate surroundings, forming bubbles that expand until their ionizing photons are consume by the neutral IGM. As the number of radiating sources increases, so do the number and size of the ionization bubbles, which eventually spread to fill the whole Universe. However, most of the details of this scenario are yet to be clarified. For example: what controls the formation of the first objects and how much ionizing radiation do they produce? How do the bubbles expand into the intergalactic medium and what do they ionize first, high-density or low density regions? The answer to these important questions and many others touch upon many fundamental questions in cosmology, galaxy formation, quasars activity and the physical properties of very metal poor stars [9, 30, 44, 41, 67, 135].

To ionize hydrogen one needs photons with energy of 13.6 eV or higher meaning the reionization of the Universe requires ultraviolet photons. A crucial question is which sources in the Universe provide the UV photons needed to ionize the Universe and maintain it in that state. Obvious candidates are the first stars (so called Population III stars), second generation stars (Population II stars) and (mini)quasars which are objects powered by intermediate mass black holes \((10^{3-6}M_\odot)\). There are other candidate sources of reionization, like decaying or self-annihilating dark matter particles or decaying cosmic strings. However, the constraints on such objects make it unlikely that they could reionize the Universe by themselves [40, 100, 121, 122, 142, 151, 168, 231].

Massive black holes powering quasars convert mass to radiation extremely efficiently. They produce a large amount of UV and X-ray radiation above the ionization threshold. In fact, one of the main discoveries of the last decade is that quasars, powered by very large black holes with masses in excess of \(10^9M_\odot\), already existed at redshift above 7 (e.g., QSO ULAS J1120+0641 [137] from the UKIDSS survey [108]). How these black holes managed to accumulate so much mass in such a
short time is a puzzle in its own right [137, 20]. However, the mass distribution of the massive black holes in the early Universe is unknown, rendering the role played by quasars during reionization very uncertain.

Population III stars formed from the primordial mix of the elements and thus only contain hydrogen and helium. This composition makes them very different from present-day stars. In order for a star to form, the initial proto-star has to radiate some of the energy gained by gravitational contraction, or the collapse will rapidly halt as the cloud reaches hydrostatic equilibrium. Population III stars are poor radiators until the cloud from which they form reaches high temperatures. This causes them to be very massive, and hence, they are very efficient and abundant sources of UV photons, yet are very short lived. Theoretically, these objects could have reionized the Universe but our knowledge of them, including the question of whether they existed in sufficient numbers, is very uncertain.

The first stars’ early metal enrichment was likely the dominant effect that brought about the transition from Population III to Population II star formation. Recent numerical simulations of collapsing primordial objects with masses of $\approx 10^6 M_\odot$, have shown that the gas has to be enriched with heavy elements to a minimum level of $Z_{\text{crit}} \approx 10^{-4} Z_\odot$, in order to have any effect on the dynamics and fragmentation properties of the system. Normal, low-mass (Population II) stars are hypothesized to form only out of gas with metallicity $Z \geq Z_{\text{crit}}$. Thus, the characteristic mass scale for star formation is expected to be a function of metallicity, with a discontinuity at $Z_{\text{crit}}$ where the mass scale changes by about two orders of magnitude. The redshift where this transition occurs has important implications for the early growth of cosmic structure, and the resulting observational signature includes the extended nature of reionization (see the review by Ciardi and Ferrara [44]).

Most studies of reionization have focused on stars as being the primary source [126, 1, 2, 29, 221]. Due to the deficiency of hard photons in the spectral energy distributions (SEDs) of these “first stars”, heating due to these objects is limited in extent [204]. On the other hand, miniquasars (miniquasos), characterized by central black hole masses $< 10^6 M_\odot$, have also been considered as an important contributor to reionization [120, 165, 166, 144, 66, 69, 217, 204]. Ionization aspects of the miniquasar radiation have been explored by several authors [120, 165, 166, 204, 206, 229]. Thomas & Zaroubi [204] have shown that although the ionization pattern around miniqsos is similar to that of stellar-type sources, the heating due to the presence of hard photons in miniqsos is very different. The reason being is that stars produce thermal radiation that is mostly in the UV range, which is very efficient in ionization, but once it is absorbed by $H_\alpha$, the energy left will be too small to be converted to heat effectively. On the other hand black hole powered sources have hard x-ray photons as their spectral energy distribution (SED) follows a power law (typically assumed to be $-1$). Such x-ray photons have lower bound-free cross section relative to UV photons but once they are absorbed, their leftover energy is very large and can easily be converted to heat. Also, x-ray photons penetrate much deeper into the IGM and can heat it up much further from the source than UV radiation.
Miniqsos heat the surrounding IGM well beyond their ionization front \cite{204, 42}. Several authors (e.g., \cite{120, 144, 230}) have shown the importance of heating the IGM with respect to the observability of the redshifted 21 cm radiation in either emission or absorption. Figure 13 shows the ionization and heating patterns around a number of stars (upper panels) and miniqsos (lower panel). The mass of the stars and black-holes are indicated next to the lines, and their SEDs are assumed to be thermal or to have a power law dependence on the photon energy, $\propto E^{-1}$, respectively. The calculation here is spherically symmetric and assumes a single object forming in the IGM \cite{230, 204}. The ionization pattern around stars and black holes are very similar, they both show a very abrupt increase in H$^+$ with a clear ionization front (see e.g. \cite{102, 204, 230}). Of course the radius at which such front is seen depends on the mass of the star or the black hole but the pattern is the same (see the left hand side panels of Figure 13). The heating profile around the two types of sources, on the other hand, is different since in power law sources (miniqsos) the radiation can penetrate the neutral gas and reach large distances from the sources (see right hand panels of Figure 13). This high energy radiation is efficient in heating the IGM gas through secondary electrons \cite{183} (see discussion later) whereas UV radiation is efficient in ionizing the gas but has little energy left to heat too much and can not penetrate the neutral gas as far as x-ray radiation does.

We have seen that unlike stars, x-ray source a (e.g., miniquasars) have an additional property of heating the IGM to a large extent and through secondary Lyman $\alpha$ radiation making the neutral IGM visible to a 21-cm experiment. However, some authors (e.g., \cite{52, 175}) argue that miniquasars alone can not reionize the Universe as they will produce far more soft X-ray background radiation than currently observed \cite{136, 186} while simultaneously satisfying the WMAP3 polarisation results \cite{152, 187}. It should be noted, however, that these calculations have been carried out assuming specific models for the evolution of black hole mass density and spectral energy distributions of UV/X-ray radiation of the miniquasars, whereas one can easily construct other models in which the discrepancy is not so severe \cite{230, 169}.

Some authors \cite{99} have claimed a detection of excess IR background radiation and argued that it provides evidence for stars being the primary source of reionization. This too has been subject to controversy because of the sensitivity of the result to the subtraction of the contaminants, e.g., Zodiacal light, within the same waveband \cite{48}.

Although uncertainty looms about the sources that resided during the *dark ages*, it is conceivable from observations of our Universe up to redshifts of 6.5, that sources of reionization could have been a mixture of both stellar and miniquasar sources. Implementing radiative transfer that includes both ionizing and hard X-ray photons has been difficult and, as a result, most 3-D radiative transfer schemes restrict themselves to ionization due to stars \cite{12, 45, 73, 126, 127, 141, 154, 170, 163, 196, 211, 222}. In \cite{165}, a “semi” hybrid model of stars and quasars like the one hinted above used, albeit in sequential order instead of a simultaneous implementation. That is, pre-ionization due to quasars has been invoked between $7 \leq z \leq 20$, after which stars reionize the Universe at redshift 7.
The Epoch of Reionization

Fig. 13  This figure shows the ionization and heating profile around a single star and black hole forming in the IGM assuming spherical symmetry. The upper panels show the stars case whereas the lower panels show the black holes case. The left hand panels show the neutral fraction of H i as a function of distance from the star and the right hand panels show the gas temperature as a function of distance from the source [204].

Given the numerical cost of the full 3D radiative transfer schemes, exploring a large parameter space for models of reionization, is not feasible. Such an exploration is needed in order to understand the various physical effects introduced by each such parameter. It is also needed to help interpret the available data. A number of authors have been pursuing “quick-and-dirty” methods to simulate the reionization process. These schemes can include very rough methods that use the initial density field to produce a reionization cube without the need for cosmological N-body and hydro simulations, such as 21cmFAST ([129][222][223]) and SimFast21 [177]. They also include more accurate (yet still fast) methods like BEARS [204][206][205] that use N-body and hydro simulations but reduces the numerical cost by restricting the ionization bubbles around the radiation sources to be spherical.
4 The redshifted 21 cm as a probe of the EoR

In recent years it has become clear that the 21 cm line can be used to probe the neutral IGM prior to and during the reionization process. This hyperfine transition line of atomic hydrogen (in the ground state) arises due to the interaction between the electron and proton spins [83, 180, 120]. The excited triplet state is a state in which the spins are parallel whereas the spins at the lower (singlet) state are antiparallel. The 21 cm line is a forbidden line for which the probability for a spontaneous $1 \rightarrow 0$ transition is given by the Einstein $A$ coefficient that has the value of $A_{10} = 2.85 \times 10^{-15} \text{sec}^{-1}$. Such an extremely small value for Einstein-$A$ corresponds to a lifetime of the triplet state of $1.1 \times 10^7$ years for spontaneous emission. Despite its low decay rate, the 21 cm transition line is one of the most important astrophysical probes, simply due to the vast amounts of hydrogen in the Universe [58, 209, 139] as well as the efficiency of collisions and Lyman-$\alpha$ radiation in pumping the line and establishing the population of the triplet state [215, 61]. In this chapter I will describe the basic physics behind this transition, especially what decides its intensity.

4.1 The 21 cm Spin and Brightness Temperatures

The intensity of the 21 cm radiation is controlled by one parameter, the so called spin temperature, $T_{\text{spin}}$. This temperature is defined through the equation,

$$\frac{n_1}{n_0} = 3 \exp(-T_\nu/T_{\text{spin}}),$$

(9)

where $n_1$ and $n_0$ are the number densities of electrons in the triplet and singlet states of the hyperfine level respectively, and $T_\nu = 0.0681 \text{ K}$ is the temperature corresponding to the 21 cm wavelength. The spin temperature is therefore, merely a shorthand for the ratio between the occupation number of the two hyperfine levels. This ratio establishes the intensity of the radiation emerging from a cloud of neutral hydrogen. Of course, in the measurement of such radiation one has to take into account the level of background being transmitted through a given cloud as well as the amount of absorption and emission within the cloud. Namely, one has to use the equation of radiative transfer.

In the following derivation I follow the description in Rybicki and Lightman (173). The radiative transfer equation is normally written in terms of the brightness (or specific intensity) of the radiation $I_\nu$. This quantity is defined as the intensity per differential frequency element in the form, $I_\nu = \frac{dI}{d\nu}$, where $\nu$ is the frequency. The intensity has the dimensions of ergs $\text{s}^{-1} \text{cm}^{-2} \text{sr}^{-1} \text{Hz}^{-1}$, namely, it quantifies the energy carried by radiation traveling along a given direction, per unit area, frequency, solid angle, and time. The radiative transfer equation for thermally emitting material at temperature $T$ can be written in terms of the optical depth for absorption
as,

\[
\frac{dI_\nu}{d\tau_\nu} = -I_\nu + B_\nu(T),
\]

(10)

where \( \tau_\nu \) is the optical depth for absorption through the cloud at a given frequency and \( B_\nu \) is the Planck function.

Fig. 14 A cartoon that shows that set up of the various components relevant to radiative transfer problem at hand starting from the background (CMB) radiation, going through a certain cloud with temperature \( T_{\text{spin}} \) and emerging with a temperature \( T_b \) that is measured by our telescopes.

In radio astronomy the intensity \( I_\nu \) is often expressed by its equivalent brightness temperature, \( T_b(\nu) \). This is convenient because at the Rayleigh-Jeans low energy limit, the relation between the brightness temperature and specific intensity is given by,

\[
T_b(\nu) \approx I_\nu c^2 / 2k_B v^2,
\]

(11)

where \( c \) is the speed of light and \( k_B \) is the Boltzmann’s constant. Expressing the radiative transfer equation [10] in terms of the brightness temperature gives it a particularly simple form,

\[
\frac{dT_b}{d\tau_\nu} = -T_b + T_{\text{CMB}},
\]

(12)

where I substituted the CMB temperature for the background temperature. Solving equation [12] yields the temperature of the emergent radiation at frequency \( \nu \),

\[
T_b(\nu) = T_{\text{spin}}(1 - e^{-\tau_\nu}) + T_{\text{CMB}}(\nu)e^{-\tau_\nu},
\]

(13)

where \( T_{\text{spin}} = T_b(0) \) is the brightness temperature in the absorbing cloud (see Figure[14]). Notice that for the background radiation the factor \( \exp(-\tau_\nu) \) gives the transmission probability of the background radiation whereas the \( 1 - \exp(-\tau_\nu) \) factor gives the emission probability of 21 cm photons from within the cloud. Therefore, in order to determine the brightness temperature, one needs to know the optical depth for absorption, \( \tau_\nu \), and the spin temperature, \( T_{\text{spin}} \), in the optically thin regime relevant to our case. Notice that in the case in which \( T_{\text{spin}} = T_{\text{CMB}} \) the brightness temperature gives exactly the CMB temperature. This is simply because in such a case there is a prefect balance between the absorption and emission at every frequency. Therefore, the measurement in such a case does not reveal anything interesting about the intervening cloud, the subject we are interested in here.
I will first start with calculating the 21 cm optical depth. The hyperfine transition of atomic hydrogen is an ideal transition to be described by Einstein coefficients and their relations. The 21 cm radiation incident on the atom can cause $0 \rightarrow 1$ transitions (absorptions) and $1 \rightarrow 0$ transitions (induced emissions) corresponding to Einstein coefficient $B_{01}$ and $B_{10}$ respectively. The probabilities are given by,

$$I_\nu B_{01} = \frac{g_1}{g_0} B_{10} I_\nu,$$

and

$$I_\nu B_{10} = A_{10} \frac{\lambda^2 I_\nu}{2h\nu_{10}},$$

respectively \cite{173}. Here $\nu_{10} = 1420.4$ MHz is the frequency of the 21 cm transition.

The 21 cm line absorption cross section is given by

$$\sigma_\nu \equiv \sigma_{01} \phi(\nu) = \frac{3e^2A_{10}}{8\pi\nu^2} \phi(\nu),$$

where $\phi(\nu)$ is the line profile defined so that $\int d\nu \phi(\nu) = 1$ and has units of time.

The optical depth of a cloud of hydrogen is then:

$$\tau_\nu = \int d\ell \sigma_{01} (1 - e^{-E_{10}/k_B T_{\text{spin}}}) \phi(\nu) n_0$$

$$\approx \sigma_{01} \left( \frac{\nu}{k_B T_{\text{spin}}} \right) \left( \frac{N_{\text{HI}}}{4} \right) \phi(\nu),$$

where $N_{\text{HI}}$ is the column density of H I and $d\ell$ is a line element within the cloud. The factor of 4 connecting $n_0$ and H I accounts for the fraction of atoms in the hyperfine singlet state. The second factor in equation (17) with $E_{10}$ accounts for stimulated emission. The approximate form in equation (18) assumes uniformity throughout the cloud.

We now substitute for $\phi(\nu)$ and $N_{\text{HI}}$ using cosmological quantities. In general, the line shape $\phi(\nu)$ includes natural, thermal, turbulent and velocity broadening, as well as bulk motion (which increases the effective Doppler spread). Velocity broadening is the most important effect in the IGM. Hubble expansion of the gas results in velocity broadening of a region of linear dimension $\ell$ will be $\Delta\nu \sim \ell H(z)$ so that $\phi(\nu) \sim c/(\ell H(z) \nu)$. The column density along such a segment depends on the neutral fraction $x_{\text{HI}}$ of hydrogen, so $N_{\text{HI}} = \ell x_{\text{HI}} n_{\text{H}}(z)$ \cite{67}. A more exact solution of equation (17) yields an expression for the 21 cm optical depth of the diffuse IGM,

$$\tau_{v_0} = \frac{3}{32\pi} \frac{h \nu^2}{k_B T_{\text{spin}} n_0^2} \frac{x_{\text{HI}} n_{\text{H}}}{(1 + z) (d\nu//dr)}$$

$$\approx 0.0092 (1 + \delta) (1 + z)^{3/2} x_{\text{HI}} \frac{T_{\text{spin}}}{n_{\text{H}}}(H(z)/(1 + z)),$$

where $\delta$ is the cosine of the angle between the line of sight and the direction of the gas motion.
where in the second relation, $T_{\text{spin}}$ is in degrees Kelvin. Here the factor $(1 + \delta)$ is the fractional overdensity of baryons and $dv / d\parallel$ is the gradient of the proper velocity along the line of sight, including both the Hubble expansion and the peculiar velocity [96]. In the second line, we have substituted the velocity $H(z)/(1 + z)$ appropriate for the uniform Hubble expansion at high redshifts.

Next we need to calculate the spin temperature and substitute in Eq. 13. In his seminal papers, George Field [61, 62], used the quasi-static approximation to calculate the spin temperature, $T_{\text{spin}}$, as a weighted average of the CMB temperature, $T_{\text{CMB}}$, the gas kinetic temperature, $T_{\text{kin}}$, and the temperature related to the existence of ambient Lyman-α photons, $T_\alpha$ [215, 62]. For almost all interesting cases, one can safely assume that $T_{\text{kin}} = T_\alpha$ [61, 67, 120, 135].

Three competing processes determine $T_{\text{spin}}$: (1) absorption of CMB photons (as well as stimulated emission); (2) collisions with other hydrogen atoms, free electrons, and protons; and (3) scattering of Lyman α photons through excitation and deexcitation. Hence, the spin temperature could be recast as [61]:

$$T_{\text{spin}} = \frac{T_{\text{CMB}} + y_{\text{kin}} T_{\text{kin}} + y_\alpha T_\alpha}{1 + y_{\text{kin}} + y_\alpha}$$

where $y_{\text{kin}}$ and $y_\alpha$ are the kinetic and Lyman-α coupling terms, respectively. It is important to note that for the 21 cm radiation to be observed, it has to attain a different temperature than that of the CMB background [61, 63, 62, 83, 215]. The form I use here for Eq. 21 is the original form used in the George Field’s 1958 paper [61], whereas some authors use a form that relates the inverse of the various temperatures. Both ways are of course equivalent but one needs to be careful with the definitions of the coupling coefficients in each case.

The kinetic coupling term $y_{\text{kin}}$ is due to collisional excitations of the 21 cm transitions. The Lyman-α coupling term $y_\alpha$ is due to the so called Lyman-α pumping mechanism, also known as the Wouthuysen-Field effect, which is produced by photo-exciting the hydrogen atoms to their Lyman transitions [61, 62, 215]. The coupling factors $y_{\text{kin}}$ and $y_\alpha$ depend on the rate of collisional and Lyman α pumping within the H I cloud. A number of authors have calculated these rates in detail [6, 114, 185, 212, 234]. In the case of first stars, the Wouthuysen-Field effect will depend on the intensity of the Lyman α photons produced by these sources. Collisions on the other hand are somewhat more complicated since it is normally done through the so called secondary electrons which are released by the ionization of an H I atom by an x-ray photon. An electron with such high energy will lose it to the rest of the IGM through collisions. This energy will in general be divided between collisional excitation, collisional ionization and heating [65, 68, 183, 208].

Since decoupling mechanisms can influence the spin temperature in different ways, it is important to explore the decoupling issue for various types of ionization sources. For instance, stars decouple the spin temperature mainly through radiative Lyman α pumping whereas x-ray sources (e.g., mini-quasars) decouple it through a combination of collisional excitation and heating [42, 230], both produced by the energetic secondary electrons ejected due to x-ray photons [183]. The difference in
the spin temperature decoupling patterns of the two, will eventually help disentangle the nature of the first ionization sources \cite{204, 159}.

Collisions could also be induced by Compton scattering of the CMB photons off the residual free electrons in the IGM gas. This process is dominant at high redshifts $z \gtrsim 200$ and keeps the gas temperature equal to that of the CMB. However, it is not efficient enough at lower redshifts to heat the gas, it is still sufficient to couple the spin temperature to the gas down to $z \approx 100$. In fact, one can show that the global spin temperature evolves in an intricate fashion bouncing back and forth between the gas (kinetic) temperature and the CMB temperature based on which heating/excitations mechanism is dominant.

Figure 15 shows the expected global evolution of the spin temperature as a function of redshift. The blue solid line represents $T_{\text{CMB}}$, which drops as $1 + z$. The green line shows the gas temperature as a function of redshift. At $z \gtrsim 200$, the gas temperature is still coupled to the CMB due to Compton scattering of the background photons off residual electrons leftover from the recombination era. At redshift $\sim 200$, however, the gas decouples from the CMB radiation and starts adiabatically cooling as a function of the redshift squared, $(1 + z)^2$, until the first objects start forming and heating up the gas at redshift below 30. The spin temperature (shown by the red lines) has a somewhat more complicated behavior. At $z \approx 100$ it is coupled to the gas temperature due to collisional coupling caused by residual electrons leftover from recombination. At $z \approx 100$ the efficiency of collisional coupling to the gas drops due to the Hubble expansion. At this stage, the spin temperature starts veering towards $T_{\text{CMB}}$ until it is completely dominated by it. At lower redshifts the first astrophysical objects that heat and ionize the IGM couple $T_{\text{spin}}$ to the gas. Here, broadly speaking, there are two possible histories, one in which $T_{\text{spin}}$ couples to the gas as it heats up once it obtains a temperature greater than $T_{\text{CMB}}$ (red solid line). In the other possible evolution the spin temperature couples to the gas much before the kinetic temperature exceeds that of the CMB (red dashed line) \cite{8, 160, 205}. In the former case the 21 cm radiation, after decoupling from the CMB at $z \lesssim 30$, is seen only in emission, whereas in the latter case it is seen initially in absorption and only at later stages in emission.

Currently all attempts to measure the redshifted 21 cm emission from the IGM are focused on the redshift range $6 \lesssim z \lesssim 12$. This is due to a number of reasons that are related to the limitations posed by the ionosphere and the background noise (see section 5 for more detail). In this range of redshifts the spin temperature is expected to be set by the astrophysics of the first objects in the Universe, namely, gas physics, feedback, etc., which often involve very complicated and poorly understood processes. However, observing the spin temperature of the Universe within the redshift window around $z \approx 50 - 100$ will mostly probe the cosmological density field \cite{116}. Such a measurement could provide a vast amount of information about the pristine Universe that, given the span of its redshift coverage, could potentially exceed that of the CMB data. Unfortunately however, the ionosphere at such frequencies $\nu \lesssim 30$ MHz poses insurmountable hurdles that render such attempts futile. This has led some authors to propose setting up radio telescopes at these very low frequencies on the moon (see e.g., \cite{109}).
The Epoch of Reionization

4.2 The Differential Brightness Temperature

As we mentioned above the measured quantity in radio astronomy is the brightness temperature, or more accurately the so called differential brightness temperature $\delta T_b = T_b - T_{\text{CMB}}$ which reflects the fact the only meaningful brightness temperature measurement insofar as the IGM is concerned is when it deviates from $T_{\text{CMB}}$. In order to get this quantity one should substitute the various components into Equation [13]. Such a substitution and rearrangement yields, [61, 62, 120, 46],

$$\delta T_b = 28\text{mK} \left(1 + \delta\right) x_{\text{HI}} \left(1 - \frac{T_{\text{CMB}}}{T_{\text{spin}}} \right) \left( \frac{\Omega_b h^2}{0.0223} \right) \sqrt{\frac{1+z}{10}} \left( \frac{0.24}{\Omega_m} \right) \left[ \frac{H(z)}{(1+z)} \right] \left( \frac{dV/\parallel}{dr/\parallel} \right),$$

(22)

where $h$ is the Hubble constant in units of 100 km s$^{-1}$ Mpc$^{-1}$, $\delta$ is the mass density contrast, $x_{\text{HI}}$ is the neutral fraction, and $\Omega_m$ and $\Omega_b$ are the mass and baryon densities in units of the critical density. Note that the three quantities, $\delta$, $x_{\text{HI}}$ and $T_{\text{spin}}$, are all functions of 3D position. The term $(T_{\text{spin}} - T_{\text{CMB}})/T_{\text{spin}}$ can obtain a maximum of +1 for $T_{\text{spin}} \gg T_{\text{CMB}}$, i.e., in the emission case. It has no such bound for the case of $T_{\text{spin}} \ll T_{\text{CMB}}$ and can be very negative in the absorption case.
Equation 22 shows that the differential brightness temperature is composed of a mixture of cosmology dependent and astrophysics dependent terms. This makes the equation a complex yet also a very informative one. This is simply because at different stages in the evolution of this field $\delta T_b$ is dominated by different contributions. For example, at high redshifts and before significant ionization takes place, i.e. $x_{\text{HI}} \approx 1$, everywhere the brightness temperature is proportional to the density fluctuations making its measurement an excellent probe of cosmology. However, at low redshifts ($z \approx 7$) a significant fraction of the Universe is expected to be ionized and the measurement is dominated by the contrast between the neutral and ionized regions, hence, probing the astrophysical source of ionization (see e.g., [90, 206]). Here I assumed that $T_{\text{spin}} \gg T_{\text{CMB}}$ at all redshifts. Figure 1, which have discussed before, shows a typical distribution of the differential brightness temperature. The figure is taken from the simulations of Thomas et al. [206]. Most radiative transfer simulations assume that the spin temperature is much larger than the CMB temperature, namely the term $\left(1 - T_{\text{CMB}}/T_{\text{spin}}\right)$ in eq. 22 is unity. As figure 15 shows, this is a good assumption at the later stages of reionization, however, it is probably not valid at the early stages. Modeling this effect is somewhat complex and requires radiative transfer codes that capture the Lyman-\(\alpha\) line formation and multifrequency effects, especially those coming from energetic photons [8, 129, 159, 205].

Here we show the evolution of the brightness temperature for three reionization histories: (1) With reionization, excitation and heating dominated by power law sources (miniqsos with x-rays); (2) dominated by thermal (stellar) sources; (3) dominated by a mixture of the aforementioned two types of sources. To create a contiguous observational cube or “frequency cube” (right ascension (RA) $\times$ declination (DEC) $\times$ redshift), the RA and DEC slices, taken from individual snapshots at different redshifts (or frequency), are stacked and interpolated smoothly to create a reionization history. This datacube is then convolved with the point spread function of the LOFAR telescope to simulate the mock data cube of the redshifted 21-cm signal as seen by LOFAR. For further details on creating this cube, refer to [206, 205].

As expected, the signatures (both visually and in terms of the r.m.s) of the three scenarios (Fig. 16) are markedly different. In the miniqso-only scenario, reionization proceeds extremely quickly and the Universe is almost completely ($x_{\text{HI}} > 0.95$) reionized by around redshift 7. The case in which stars are the only source sees reionization end at a redshift of 6. Also in this case, compared to the previous one, reionization proceeds in a rather gradual manner. The hybrid model, as explained previously, is in between the previous two scenarios.

In the models shown here, the transition from the absorption dominated brightness temperature to the emission dominated one occurs at relatively low redshifts. The transition redshift depends sensitively on the assumptions made in each case. Other authors have explored such effects and conclude that the transition occurs at much higher redshifts (see e.g., the models in [129, 177]).

The $\delta T_b$ in Fig. 16 is calculated based on the effectiveness of the radiation flux, produced by the source, in decoupling the CMB temperature ($T_{\text{CMB}}$) from the spin
temperature ($T_s$). This flux, both in spatial extent and amplitude, is obviously much larger in the case of miniqso compared to that of stars, resulting in a markedly higher brightness temperature in both the miniqso-only and hybrid models when compared to that of the stars. However, we know that stars themselves produce Lyman $\alpha$ radiation in their spectrum. Apart from providing sufficient Lyman $\alpha$ flux to their immediate surroundings, this radiation builds up as the Universe evolves into a strong background [46], potentially filling the Universe with enough Lyman $\alpha$ photons to couple the spin temperature to the kinetic temperature everywhere. It has to be noted that the results we are discussing here are extremely model dependent and any changes to the parameters can influence the results significantly.

### 4.3 The 21 cm forest at high $z$

Finally, I will conclude this section by discussing a very different aspect of the redshifted 21 cm radiation, and that is the case of the 21 cm forest. Very bright
radio sources might exist at high redshifts. In such a case, the emission from these sources is expected to be resonantly absorbed by the neutral IGM and form a system of absorption features just like the Lyman $\alpha$ forest seen in the spectra of distant quasars. Such absorption features are called the 21 cm forest and they were first investigated by Carilli et al. [37] and subsequently by other authors [36, 66, 64, 119, 219]. The discovery of such systems will provide very valuable information about the reionization process and the IGM’s physical properties during the EoR which will be largely independent of calibration errors (see section 5). Currently, we know of no very bright high redshift sources, but with the imminent availability of highly sensitive radio telescopes like LOFAR and SKA, the prospects for detecting such sources are very promising.

Figure [17] shows a simulated spectrum at 1 kHz resolution of a $z = 10$ radio source with a flux density of 20 mJy at an observing frequency of 120 MHz ($S_{120}$). The implied luminosity density at a rest frame frequency of 151 MHz is then $P_{151} = 2.5 \times 10^{35} \text{ erg s}^{-1} \text{ Hz}^{-1}$. The left hand panel of Figure [17] shows a spectrum covering a large frequency range (100 MHz to 200 MHz, or HI 21cm redshifts from 13 to 6), whereas the right hand panel shows an expanded view of the frequency range corresponding to the HI 21cm line at the source redshift (129 MHz). At 129 MHz the spectrum shows a 1% drop due to the diffuse neutral IGM. See reference [37] for detail.

**Fig. 17**  **Left hand panel:** A simulated spectrum from 100 MHz to 200 MHz of a source with $S_{120} = 20$ mJy at $z = 10$ using the Cygnus A spectral model and assuming HI 21cm absorption by the IGM. Thermal noise has been added using the specifications of the SKA and assuming 10 days integration with 1 kHz wide spectral channels. **Right hand panel:** The same as the left panel, but showing an expanded view of the spectral region around the frequency corresponding to the redshift HI 21cm line at the source redshift (129 MHz). The solid line is the Cygnus A model spectrum without noise or absorption. Figure taken from [37].
5 The redshifted 21 cm Observation

In section 4 we discussed the cosmological 21 cm signal and showed that it is expected to be on the order of \( \approx 10 \) mK. However, the detectable signal in the frequency range that corresponds to the epoch of reionization is composed of a number of components each with its own physical origin and statistical properties. These components are: (1) the 21 cm signal coming from the high redshift Universe, (2) galactic and extra-galactic foregrounds, (3) ionospheric influences, (4) telescope response effects, (5) Radio frequency interference (RFI) \[146, 147\] and (6) thermal noise (see Figure 2). Obviously, the challenge of the experiments in the low frequency regime is to distill the cosmological signal out of this complicated mixture of influences. This will depend crucially on the ability to calibrate the data very accurately so as to correct for the ever changing ionospheric effects and variation of the instrument response with time.

Currently, there are two types of redshifted 21 cm experiments that are attempting to observe the EoR. The first type are experiments that measure the global (mean) radio signal at the frequency range of \( \nu = [100 - 200] \) MHz averaged over the whole sky (hemisphere) as a function of frequency. In this radio signal the 21 cm radiation from the EoR is hidden. The expected measurement should show an increase of the intensity at higher redshifts due to the increase in the neutral fraction of H\textsc{i}. In particular, if the reionization process occurred rapidly such a measurement should exhibit a step-like jump in the mean brightness temperature at the redshift of reionization \( z_i \) (in this case \( z_i \) is well defined). This type of measurement is cheap and relatively easy to perform. However, given the amount of foreground contamination, especially from our Galaxy, radio frequency interference (RFI), noise and calibration errors as well as the limited amount of information contained in the data (mean intensity as a function of redshift), such experiments are in reality very hard to perform. One of those experiments, EDGES \[28\], has recently reported a lower limit on the duration of the reionization process to be \( \Delta z > 0.06 \), thus providing a very weak constraint on reionization as most realistic simulations predict that this process occurs over a much larger span of redshift \[27\].

The second type of experiment is interferometric experiments carried out in the frequency range of \( \approx 100 - 200 \) MHz, corresponding to a redshift range of \( \approx 6 - 12 \). This type of experiment is considered more promising. The main reason for this is that these experiments allow better control of what is being measured and contain a huge amount of information so as to allow a much more accurate calibration of the instrument. Furthermore, radio interferometers are more diverse instruments that can be used to study many scientific topics besides the EoR, which makes them appealing for a much wider community. Having said that, however, one should note that the cost involved in building and running such facilities is much higher than for the global signal experiments.
5.1 Radio Interferometry and the Calibration Problem

Interferometers measure the spatial correlation of the electric field vector emanating from a distant source in the sky, \( E(R,t) \), located at position \( R \) and measured at time \( t \). The sketch presented in the left hand panel of Figure 18 shows the basic principle of interferometry. The two stations (dishes) receive a wavefront from a distant source and the receivers are timed to account for the difference in the pathway to the two stations which obviously depends on the source location on the sky. The signals measured at the two stations, taken with the appropriate time difference \( \Delta t \), are then cross-correlated (see for example [200, 207]).

The measured spatial correlation of the electric field between two interferometric elements (stations) \( i \) and \( j \) is called the “visibility” and is approximately given by [200, 207]:

\[
V(\nu, u, v) = \int A(l, m; \nu)I(\nu, l, m)e^{i(u l + v m)}dl dm, \tag{23}
\]

where \( A \) is the normalized station response pattern and \( I \) is the observed intensity at frequency \( \nu \). The coordinates \( l \) and \( m \) are the projections (direction cosines) of the source in terms of the baseline in units of wavelength. As a side note, here we ignore the effect of the Earth’s curvature, the so called \( w \)-projection. From this equation it is clear that the observed visibility is basically the Fourier transform of the intensity measured at the coordinates \( u \) and \( v \). Notice that coordinates \( u \) and \( v \) depend on the baseline and its direction relative to the source position (see the right hand panel of Figure 18). Therefore, the coordinates \( u \) and \( v \) produced by a given baseline vary with time due to Earth’s rotation and will create an arc in the uv plane that completes half of the drawn track after 12 hours as seen in the right hand panel of Figure 18. The other half is produced due the fact that the intensity is a real function. The width
of the uv-track is set by the size of the station, such that large stations produce thick tracks. I will discuss the issue of uv coverage in more detail below. One also should note that the coordinates u and v are a function of wavelength, namely their value will change as a function of frequency, which one has to take into account when combining or comparing results from different frequencies.

In the interferometric visibilities there always exist errors introduced by the sky, the atmosphere (e.g. troposphere and ionosphere), the instrument (e.g. beam-shape, frequency response, receiver gains etc.) and by Radio Frequency Interference (RFI). The process of estimating and reducing the errors in these measurements is called calibration and is an essential step before understanding the measured data. Calibration normally involves knowing very well the position and intensities of the bright sources within and without the field of view of the radio telescope and using them to correct for the ionospheric and instrumental effects introduced into the data [76,101,155,220]. This is similar to the adaptive optics techniques used in the optical regime except that here one needs to account for the variations in polarization of the radiation as well as in its total intensity.

Since most current instruments are composed of simple dipoles as their fundamental elements which have a polarized response (preferred x and y direction), the main danger in insufficient calibration lies in the possible leakage of polarized components into total intensity, thereby severely polluting the signal (see e.g., [94]). That is to say, since the cosmological signal is not expected to be polarized, if the polarized response of the instrument is not very well understood and taken into account it will mix some of the polarization that exists in the Galactic foregrounds (see subsection 5.5) with the cosmological signal and create a spurious signal that can not be distinguished from the cosmological signal. Hence, a very accurate calibration of these instruments is absolutely needed. Another issue one needs to deal with is that of the Radio Frequency interference, but we will not discuss it here and refer the reader instead to the papers by Offringa et. al. [146, 147].

5.2 Current and Future EoR Experiments

Currently, there are a number of new generation radio telescopes, GMRT\textsuperscript{2}, LOFAR\textsuperscript{3}, MWA\textsuperscript{4}, 21CMA\textsuperscript{5} and PAPER\textsuperscript{6} that plan to capture the lower redshift part of the $\delta T_b$ evolution ($z \approx 12$). Unfortunately, however, none of these experiments has enough signal-to-noise to provide images of the EoR as it evolves with redshift. Instead, they are all designed to detect the signal statistically. In what follows I will focus on LOFAR more than the other telescopes, simply because this is the instru-
The LOw Frequency ARray (LOFAR) is a European telescope built mostly in the Netherlands and has two observational bands, a low band and a high band covering the frequency range of 30-85 MHz and 115-230 MHz, respectively. The high band array is expected to be sensitive enough to measure the redshifted 21 cm radiation coming from the neutral IGM within the redshift range of \(z=11.4\) (115 MHz) to \(z=6\) (203 MHz), with a resolution of 3-4 arcminutes and a typical field of view of \(\sim 120\) square degrees (with 5 beams) and a sensitivity on the order of 80 mK per resolution element for a 1 MHz frequency bandwidth. At frequencies below the FM band, probed by the low band array, the LOFAR sensitivity drops significantly and the sky noise increases so dramatically (roughly like \(\approx \nu^{-2.6}\)) that detection of HI signals at these frequencies is beyond the reach of LOFAR [77, 95, 103] and all other current generation telescopes for that matter. Figure 19 shows an artistic impression of the LOFAR telescope and its spread over Europe (left hand panel). The right hand panel shows the inner most center of the core located in the north of the Netherlands and shows the two types of stations used in the array.

In the future, SKA\[^{7}\] will significantly improve on the current instruments in two major ways. Firstly, it will have at least an order of magnitude higher signal-to-noise which will allow us to actually image the reionization process. It will also give us access to the Universe’s dark ages up to redshifts as high as \(z \approx 30\) (assuming lowest frequency of about 50 MHz), hence, providing crucial information about cosmology which none of the current telescopes is able to probe. Thirdly, SKA will have a resolution better by a factor of few, at least, relative to the current telescopes [228]. These three advantages – sensitivity, resolution and frequency coverage – will not only improve on the understanding we gain with current telescopes but give the opportunity to address a host of fundamental issues that current telescopes will not be able to address at all. Here, I give a few examples: (1) due to the limited resolution and poor signal-to-noise of current telescopes, the nature of the ionizing sources is expected to remain poorly constrained; (2) the mixing between the astrophysical effects and the cosmological evolution is severe during the EoR but much less so during the dark ages, an epoch beyond the reach of the current generation of telescopes, but within SKA’s reach; (3) at redshifts larger than 30, the 21 cm could potentially provide very strong constraints, potentially much more so than the CMB, on the primordial non-gaussianity of the cosmological density field, which is essential in order to distinguish between theories of the very early Universe (e.g., between different inflationary models).

\[^{7}\] Square Kilometer Array, \text{http://www.skatelescope.org/}
Fig. 19 Left hand panel: An artist’s impression of the layout of the LOFAR telescope over Western Europe. For the EoR, only the central part of the telescope is relevant. (courtesy of Peter Prijs). Right hand panel: The very central area of LOFAR. This circular area is known as the superterp and is the heart of the LOFAR core. The high-band array stations (covered in black blastic sheets) are clearly seen in this picture. In between one can also see the Low-Band Array antennas.

5.3 Station configuration and uv coverage

In principle, Fourier space measurement and real space measurement are equivalent. However, this is only true if one has a perfect coverage of both spaces. In reality, each baseline will cover a certain line in the so-called uv plane which needs to be convolved with the width of the track (see right hand panel of Figure 18). The combination of all the tracks of the array produces the uv coverage of the interferometer. The low frequency arrays must be configured so that they have a very good uv coverage. This is crucial to the calibration effort of the data where a filled uv plane is important for obtaining precise Local \([143]\) and Global \([184]\) Sky models (LSM/GSM; i.e. catalogues of the brightest, mostly compact, sources in and outside of the beam, i.e. local versus global). It is also crucial for the ability to accurately fit for the foregrounds \([78, 95]\) and to the measurement of the EoR signal power spectrum \([25, 77, 82, 176]\).

The uv coverage of an interferometric array depends on the layout of the stations (interferometric elements), their number and size as well as on the integration time, especially, when the number of stations is not large enough to have a good instantaneous uv coverage.

For a given total collecting area one can achieve a better uv coverage by having smaller elements (stations). For example LOFAR has chosen to have large stations resulting in about \(\approx 10^3\) baselines in the core area. Such a small number of baselines needs about 5-6 hours of integration time per field in order to fill the uv plane (using the Earth’s rotation). In comparison, MWA which has roughly 1/3 of the total collecting area of LOFAR but chose to have smaller stations with about \(\approx 10^5\) baselines resulting in an almost instantaneous full uv coverage.

The decision on which strategy to follow has to do with a number of considerations that include the ability to store the raw visibilities, hence, allowing for a better calibration and an acceptable noise level for both the foreground extraction
Saleem Zaroubi

needs as well the power spectrum measurement (see the following sections; sec. 5.4 and sec. 5.5). A compromise between these issues as well the use of the telescopes for scientific projects other than the EoR is what drives the decision on the specific layout of the antennas.

5.4 Noise Issues

Fig. 20 This figure shows how two different experiments might sample an annulus in the uv plane. The size of uv point is given by the station (interferometric element) size where a larger station (left hand panel) has a larger footprint relative to the smaller station case (right hand panel) in the uv plane; the footprint is shown by the purple circles. Even though the sampled area in the two cases might be the same, the fact that smaller stations sample the annulus more results in an increased accuracy in their estimation of the power spectrum.

In the low frequency regime the random component of the noise, i.e., the thermal noise, is set by two effects: the sky noise and the receiver noise. At frequencies $\nu$ below $\approx 160$MHz the sky is so bright that the dominant source of noise is the sky itself, whereas at higher frequencies the receiver noise starts to be more important. The combination of the two effects is normally written in terms of the so called system temperature, $T_{\text{sys}}$. One can show that the thermal noise level for a given visibility, i.e., uv point, is,

$$\Delta V(u,v) \approx \frac{2k_B T_{\text{sys}}}{\epsilon dA \sqrt{Bt}},$$  \hspace{2cm} (24)$$

where $\epsilon$ is the efficiency of the system, $dA$ is the station area, $B$ is the bandwidth and $t$ is the observation time (see e.g., [133]). This expression is simple to understand in that the more one observes – either in terms of integration time, frequency
bandwidth or station collecting area – the less uncertainty one has. Obviously, if the signal we are after is well localized in either time, space or frequency the relevant noise calculation should take that into account.

In order to calculate the noise in the 3D power spectrum, the main quantity we are after, one should remember that the frequency direction in the observed datacube can be mapped one-to-one with the redshift, which in turn can be easily translated to distance, whereas the \( u \) and \( v \) coordinates are in effect Fourier space coordinates. Therefore, to calculate the power spectrum one should first Fourier transform the data cube along the frequency direction. Following Morales’ work [133] I will call the new Fourier space coordinate \( \eta \) (with \( d\eta \) resolution), which together with \( u \) and \( v \) defines the Fourier space vector \( \mathbf{u} = \{u, v, \eta\} \). From this, one can calculate the noise contribution to the power spectrum at a given \(|\mathbf{u}|\),

\[
P_{\text{noise}}(|\mathbf{u}|) \approx 2N_{\text{beam}}^{-1}N_{\text{cell}}^{-1/2} \left( \frac{2k_B T_{\text{sys}}}{e\Delta A d\eta} \right)^2 \frac{1}{Bn(|\mathbf{u}|)\eta},
\]

where \( N_{\text{beam}} \) is the number of simultaneous beams that could be measured, \( N_{\text{cell}} \) is the number of independent Fourier samplings per annulus and \( n(|\mathbf{u}|) \) is the number of baselines covering this annulus [133]. Note that \( n(|\mathbf{u}|) \) is proportional to the square of the number of stations, hence, \( n(|\mathbf{u}|)dA^2 \) is proportional to the square of the total collecting area of the array regardless of the station size. This means that in rough terms the noise power spectrum measurement does not depend only on the total collecting area, bandwidth and integration time, it also depends the number of stations per annulus. This is easy to understand as follows. The power in a certain Fourier space annulus is given by the variance of the measured visibilities in the annulus which carries uncertainty proportional to the inverse square root of the number of points. This point is demonstrated in Figure 20 [133, 134, 228].

5.5 The Foregrounds

The foregrounds in the frequency regime (40 – 200MHz) are very bright and dominate the sky. In fact the amplitude of the foreground contribution, \( T_{\text{sky}} \), at 150MHz is about 4 orders of magnitude larger than that of the expected signal. However, since we are considering radio interferometers the important part of the foregrounds is that of the fluctuations and not the mean signal, which reduces the ratio between them and the cosmological signal to about 2-3 orders of magnitude, which is still a formidable obstacle to surmount.

The most prominent foreground is the synchrotron emission from relativistic electrons in the Galaxy: this source of contamination contributes about 75% of the foregrounds. Other sources that contribute to the foregrounds are radio galaxies, galaxy clusters, resolved supernovae remnants and free-free emission, which together provide 25% of the foreground contribution (see e.g., [182]). Figure 21 shows the simulated foreground contribution at 120 MHz taking into account all the foreground sources mentioned.
Fig. 21 A figure showing the various cosmological and galactic components that contribute to the measured signal at a given frequency. The slices are color coded with different color tables owing to the vast difference in the range of brightness temperature in each component. The figure also shows the rms of the galactic foregrounds, extra galactic foregrounds and cosmological signal.

Observationally, the regime of frequencies relevant to the EoR is obviously not very well explored. There are several all-sky maps of the total Galactic diffuse radio emission at different frequencies and angular resolutions. The 150 MHz map of Landecker & Wielebinski [104] is the only all-sky map in the frequency range (100 – 200 MHz) relevant for the EoR experiments, but has only 5° resolution.

In addition to current all-sky maps, a number of recent dedicated observations have given estimates of Galactic foregrounds in small selected areas. For example, Ali et al. [5] have used 153 MHz observations with GMRT to characterize the visibility correlation function of the foregrounds. Rogers and Bowman [171] have measured the spectral index of the diffuse radio background between 100 and 200 MHz. Pen et al. [158] have set an upper limit to the diffuse polarized Galactic emission.

Recently, a comprehensive program was initiated by the LOFAR-EoR collaboration to directly measure the properties of the Galactic radio emission in the frequency range relevant for the EoR experiments. The observations were carried out using the Low Frequency Front Ends (LFFE) on the WSRT radio telescope. Three different fields were observed. The first field was a highly polarized region known as the Fan region in the 2nd Galactic quadrant at a low Galactic latitude of ~ 10°
This field is not ideal for measuring the EoR but it is a good field to learn from about calibration issues and about the influence of strong polarization.

The second field was a very cold region in the Galactic halo ($l \sim 170^\circ$) around the bright radio quasar 3C196, and the third was a region around the North Celestial Pole (NCP) \cite{14}. In the Fan region fluctuations of the Galactic diffuse emission were detected at 150 MHz for the first time (see Fig. 22). The fluctuations were detected both in total and polarized intensity, with an $rms$ of 14 K (13 arcmin resolution) and 7.2 K (4 arcmin resolution) respectively \cite{13}. Their spatial structure appeared to have a power law behavior with a slope of $-2.2 \pm 0.3$ in total intensity and $-1.65 \pm 0.15$ in polarized intensity (see Fig. 22). Note that, due to its strong polarized emission, the “Fan region” is not a representative part of the high Galactic latitude sky.

The foregrounds in the context of the EoR measurements have been studied theoretically by various authors \cite{182, 51, 50, 47, 176, 95, 71, 213, 49, 192, 210, 191, 26}. The first comprehensive simulation of the EoR foregrounds was carried out by Jelić et al. \cite{95} focusing on the LOFAR-EoR project. The Jelić et al. model takes into account the Galactic diffuse synchrotron & free-free emission, synchrotron emission from Galactic supernova remnants and extragalactic emission from radio galaxies and clusters, both in total intensity and polarization. The simulated foreground maps, in their angular and frequency characteristics, are similar to the maps expected from the LOFAR-EoR experiment (see Fig. 21).

One major problem faced when considering the LOFAR-EoR data is disentangling the desired cosmological signal from the foreground signals. Even though the foregrounds are very prominent they are very smooth along the frequency di-
rection \cite{182, 95, 94, 13, 14}, as opposed to the cosmological signal that fluctuates along the same direction. Hence, the separation of the two is, at a first glance, very simple. One would fit a smooth function to the data along the frequency direction and subtract it to obtain the desired signal. In reality, however, things are much more complicated as the existence of thermal noise and systematic errors due to calibration imperfections make the extraction much harder. In addition, the foregrounds are partially polarized, with a complicated structure along the frequency direction. The confluence of this with the ionospheric distortions and the polarized instrumental response makes it imperative to calibrate the data very accurately over a very wide field in order to obtain a very high dynamic range of observations. These factors make the fitting non-trivial, that might result in either under-fitting or over-fitting the signal. In the former case the deduced signal retains a large contribution of the foregrounds and produce a spurious “signal”. Whereas in the over-fitting case one fits out the foregrounds and some of the signal resulting in an underestimation of the cosmological signal.

The simplest method for foreground removal in total intensity as a function of frequency is a polynomial fitting performed on the log-log scale which reduces to a power law in the first order case \cite{95}. However, one should be careful in choosing the order of the polynomial to perform the fitting. If the order of the polynomial is too small, the foregrounds will be under-fitted and the EoR signal could be dominated and corrupted by the fitting residuals, while if the order of the polynomial is too big, the EoR signal could be fitted out. Arguably, it would be better to fit the foregrounds non-parametrically, i.e., allowing the data to determine their shape rather than selecting some functional form in advance and then fitting its parameters (see \cite{79}).

After foreground subtraction from the EoR observations, the residuals will be dominated by instrumental noise, i.e., the level of the noise is expected to be an order of magnitude larger than the EoR signal (assuming 300 hours of observation with LOFAR). Thus, general statistical properties of the noise should be determined and used to statistically detect the cosmological 21 cm signal, e.g., the variance of the EoR signal over the image, $\sigma_{\text{EoR}}^2$, as a function of frequency (redshift) obtained by subtracting the variance of the noise, $\sigma_{\text{noise}}^2$, from that of the residuals, $\sigma_{\text{residuals}}^2$. It has been shown the such statistical detection of the EoR signal using the fiducial model of the LOFAR-EoR experiment is possible \cite{95} (see Figure 23). Similar results by using different statistics are the skewness of the one-point distribution of brightness temperature of the EoR signal, measured as a function of observed frequency \cite{78}, and the power spectrum of variations in the intensity of redshifted 21 cm radiation from the EoR \cite{77}.
Fig. 23 This figure shows the ability to statistically extract the EoR signal from the foregrounds. Please notice the difference in the vertical axis units between the two panels. Left hand panel: One line of sight (one pixel along frequency) of the LOFAR-EoR data maps (black solid line), smooth component of the foregrounds (dashed black line), fitted foregrounds (dashed cyan line) and residuals (red solid line) after taking out of the foregrounds. Right hand panel: Statistical detection of the EoR signal from the LOFAR-EoR data maps that include diffuse components of the foregrounds and realistic instrumental noise ($\sigma_{\text{noise}}(150 \, \text{MHz}) = 52 \, \text{mK}$). Black dashed line represents standard deviation ($\sigma$) of the noise as a function of frequency, cyan dashed line is the $\sigma$ of the residuals after taking out smooth component due to the foregrounds and the red solid line the $\sigma$ of the original EoR signal. The grey shaded surface represents the 90% of detected EoR signals from 1000 independent realisations of the noise, while the cyan dashed line is the mean of the detected EoR signal. Note that the y-axis is in logarithmic scale [95].

### 6 The Statistics of the observed cosmological signal

#### 6.1 The 21 cm Power Spectrum

One of the main goals of the EoR projects is to measure the power spectrum of variations in the intensity of redshifted 21 cm radiation from the EoR [10, 15, 134, 133, 226]. As shown in Equation 22, the power spectrum depends on a number of astrophysical and cosmological quantities. The sensitivity with which this power spectrum can be estimated depends on the level of thermal noise (Eq. 25), sample variance and systematic errors arising from the extraction process, in particular from the subtraction of foreground contamination. In the LOFAR case [78, 77, 95, 94, 103], for example, the extraction process is modeled using realistic simulations of the cosmological signal, the foregrounds and the noise. In doing so we estimate the sensitivity of the LOFAR EoR experiment to the redshifted 21 cm power spectrum. Detection of emission from the EoR should be possible within 300 hours of observation with a single station beam. Integrating for longer, and synthesizing multiple station beams within the primary (tile) beam, will then enable us to extract progressively more accurate estimates of the power at a greater range of scales and redshifts (see Figure 24 taken from [77]).
6.2 High order statistics

Given the nature of the reionization process the expected signal is non-Gaussian, hence using high order statistics to characterize the data can reveal information that the power spectrum does not include. The left hand panel of Figure 25 shows the Probability Density Function (PDF) of the brightness temperature at four different redshifts; the PDF is clearly non Gaussian in all four cases. Therefore, high order
moments, like the skewness, as a function of redshift could be a useful tool for signal extraction in the presence of realistic overall levels of foregrounds and noise. Harker et al. [79], (see also [72, 89]) has shown that the cosmological signal, under generic assumptions, has a very characteristic pattern in the skewness as a function of redshift (the right hand panel of Figure 25). At sufficiently high redshifts the signal is controlled by the cosmological density fluctuations which, in the linear regime, are Gaussian. At lower redshifts, and as nonlinearity becomes important, the signal starts getting a slightly positive skewness. As the ionization bubbles begin to show up the skewness starts veering towards 0 until it crosses to the negative side when the weight of the ionized bubbles becomes more important than the high density outliers –note high density outliers are likely to ionize first– but the distribution is still dominated by the density fluctuations. At lower redshifts the bubbles dominate the PDF and the neutral areas become the “new” outliers giving rise to a sharp positive peak to the skewness. At redshift around 6 the instrument noise, assumed to be Gaussian, dominates driving the skewness again towards zero. Exploiting this characteristic behavior might allow us to pick up the cosmological signal with this high order statistic.

Fig. 25  Left hand panel: The distribution of \( \delta T_b \) in a certain cosmological simulation of reionization [90] at four different redshifts, showing how the PDF evolves as reionization proceeds. Note that the y-axis scale in the top two panels is different from that in the bottom two panels. The delta-function at \( \delta T_b = 0 \) grows throughout this period while the rest of the distribution retains a similar shape. The bar for the first bin in the bottom-right hand panel has been cut off; approximately 58 per cent of points are in the first bin at \( z = 7.78 \) [79]. Right hand panel: Skewness of the fitting residuals from data cubes with uncorrelated noise, but in which the residual image has been denoised by smoothing at each frequency before calculating the skewness. The three lines correspond to results from three different simulations [206, 89]. Each line has been smoothed with a moving average (boxcar) filter with a span of nine points. The grey, shaded area shows the errors, estimated using 100 realizations of the noise (see [79]).
6.3 Cross-correlating the LOFAR-EoR data with other data sets

Given the challenges and uncertainties involved in measuring the redshifted 21cm signal from the EoR, it is vital to corroborate this result with other probes of the EoR. Namely, other astrophysical data that probe the EoR signal. The list of such data is long but here we focus on the most promising two such probes: CMB maps and high redshift galaxy catalogs.

The CMB photons are scattered by the free electrons released during the reionization process. This scattering produces anti-correlation between the CMB signal and the EoR through a number of physical processes. This has been studied in recent years by a number of groups [7, 3, 47, 93, 174, 197, 198, 199]. Figure 26 shows a slice through reionization history of the 21cm signal (top panel) and the so called kinematic Sunyaev-Zeldovich effect which reflects the effect of reionization on the CMB photons (bottom panel) as time evolves [148, 195, 193, 194]. Notice the clear anti-correlation between the two maps. However, whereas the 21cm data store the redshift information in them the actual CMB data do not, instead they are sensitive to an integral of the map in the bottom panel over time (redshift). In both cases many instrumental, foreground and background effects might hamper the anti-correlation analysis.

Of special interest here are the CMB data expected to be collected by the Planck satellite, which will have a resolution comparable to the LOFAR core and, since it is an all-sky survey, will probe the same regions as the LOFAR-EoR project.

The other interesting data will come from high-z galaxies and quasars. These objects harbor the sources of ionization and are located at the center of the ionization bubbles. Hence, they are expected to anti-correlate with the EoR signal. Currently, a large amount of effort is being put into gathering possible galaxy candidates at high redshifts, e.g., Lyman-α emitters and Z, Y and J-bands drop outs, etc. Recent studies have shown that the anti-correlation signal will be detectable provided the number of galaxies in the regions with EoR observations is significant (e.g. see [113]).

7 Summary

The EoR, which starts about 400 million years after the Big-Bang, represents a major phase transition for hydrogen. Due to the formation of the first astrophysical sources, hydrogen in this epoch transforms from fully neutral to fully ionized. The EoR could be traced in space and time using relic radio emission that will be observed by the LOFAR radio telescope starting from the end of this year.

The EoR is determined by how and when the Universe started forming astrophysical objects and how the ionizing radiation from these objects permeates and fills the intergalactic medium. The EoR is related to many fundamental questions in cosmology, properties of the first (mini-)quasars, formation of very metal-poor stars and a slew of other important research topics in astrophysics. Hence uncovering it
The Epoch of Reionization

**Fig. 26** A slice through a simulated reionization history of the cosmological 21cm signal (top panel) and the so-called kinematic Sunyaev-Zeldovich effect which reflects the effect of reionization on the CMB photons (bottom panel) as time evolves [148, 195, 193, 194]. The expected anti-correlation between the two phenomena is patently clear. Unfortunately however, the CMB data stores no redshift (time) information and the actual map that one gets is the integral of the lower panel over redshift (time), which will make the anti-correlation more challenging to detect. This Figure is taken from [93].

will have far reaching implications on the study of structure formation in the early Universe.

Currently, there are only few observational constraints on the epoch of reionization. The CMB temperature and polarization data obtained by the WMAP satellite allow measurement of the total Thomson scattering of the primordial CMB photons off intervening free electrons produced by the epoch of reionization along the line of sight. They show that the CMB intensity has only been damped by $\sim 9\%$, indicating that the Universe was mostly neutral for 400 million years and then ionized. However, the Thomson scattering measurement is an integral constraint telling us little about the sources of reionization, its duration or how it propagated to fill the whole Universe.

Another constraint comes from specific features in the spectra of distant quasars, known as the Lyman $\alpha$ forest. These features, which are due to neutral hydrogen, indicate two important facts about reionization. First, hydrogen in the recent Universe is highly ionized, only 1 part in 10000 being neutral. Second, the neutral fraction of hydrogen in the distant Universe suddenly increases at redshift 6.5, i.e., about 900 million years after the Big Bang, demarcating the end of the reionization process.
Despite these data providing strong constraints on the ionization state of the Universe at redshifts below 6.5, they say very little about the reionization process itself. Another couple of constraints come also from the Lyman $\alpha$ forest systems, IGM temperature and the number of ionizing photons per baryon, suggesting the bulk of the reionization process occurs at late redshifts $z \approx [6 - 9]$.

A whole slew of possible constraints currently discussed in the literature are either very controversial, very weak or, as is often the case, both. Most are very interesting and exciting, but can be investigated reliably only with a new generation of instruments such as the James Webb Space Telescope, replacing the Hubble Space Telescope in the next decade.

The imminent availability of observations of redshifted 21 cm radiation from the Universe’s dark ages and the EoR will be one of the most exciting developments in the study of cosmology and galaxy and structure formation in recent years. Currently, there are a number of instruments that are designed to measure this radiation. In this contribution I have argued that despite the many difficulties that face such measurements they will provide a major breakthrough in our understanding of this crucial epoch. In particular current radio telescopes, such as LOFAR, will be able to provide us with the global history of the EoR progression, the fluctuations power spectrum during the EoR, etc., up to $z \approx 11$. These measurements will usher the study of the high redshift Universe into a new era which will bridge, at least in part, the large gap that currently exists in observation between the very high redshift Universe ($z \approx 1100$) as probed by the CMB and the low redshift Universe ($z \lesssim 6$).

Although the current generation of telescopes have a great promise they will also have limitations. For example they have neither the resolution, the sensitivity nor the frequency coverage to address many fundamental issues, like the nature of the first sources. Crucially, they will not provide a lot of information about the dark ages which is only accessible through very low frequencies in the range of $40 - 120$ ($z \approx 35 - 11$). Fortunately, in the future SKA can improve dramatically on the current instruments in terms of sensitivity, redshift coverage and resolution.

The next decade will be extremely exciting for studying the high redshift Universe, especially as these radio telescopes gradually come online, starting with LOFAR, GMRT and MWA. They promise to resolve many of the puzzles we have today pertaining to the formation and evolution of the first objects cosmology, and the physical processes in the high redshift intergalactic medium.

Acknowledgements. I would like to thank Geraint Harker, Stephen Rafter and Rajat M. Thomas for careful reading of the manuscript. Many of the results shown here have been obtained in collaboration with the members of the LOFAR EoR project whose contribution I would like to acknowledge. I would like also thank the editors of this book for giving me the opportunity to write this chapter.
References


52 Saleem Zaroubi

The Epoch of Reionization


58 Saleem Zaroubi


