The Cosmic Dawn and the Epoch of Reionization

J. M. Diego[1], C. Hernández-Monteagudo[2], D. Herranz[1], E. Martínez-González[1] and J. A. Rubiño-Martín[3]

[1] Instituto de Física de Cantabria (CSIC-UC), Av. los Castros, s/n, 39005 Santander, Spain
[2] Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan, 1, Planta 2, 44001 Teruel, Spain
[3] Instituto de Astrofísica de Canarias, 38200 La Laguna, Tenerife, Canary Islands, Spain

Abstract

In this chapter we briefly describe the physics behind the HI 21 cm line in terms of the interplay of the HI gas with the ionized plasma and the Cosmic Microwave Background, and the different phases the system undergoes as the ambient density and UV background evolve with cosmological time. We also address the problematics associated to the metal enrichment of the IGM and the implications for the models of galaxy formation and evolution. We briefly discuss possible synergies with other reionization probes like E-ELT, JWST and ALMA, and conclude by listing a number of cosmological scenarios describing some type of energetic injection in the Universe, scenarios to whose understanding SKA should be able to (at least partially) contribute.

1 Introduction

The so-called 'Cosmic Dawn', i.e. the age during which the very first sources of light (stars and galaxlies) kindled in the Universe, and the subsequent epoch of reinozation (EoR) during which most of the hydrogen of the intergalactic medium (IGM) returned to its ionized state, are not only two of the most fascinating and poorly understood phases of the evolution of the Universe, but also correspond to the range of redshifts that will be most effectively probed by the SKA. This is not an accidental coincidence. The SKA is being designed with the unveiling of the Cosmic Dawn and the EoR as one of its primary scientific goals.

In this chapter we will briefly describe the physics behind these interesting epochs. We first focus on the different states of the HI gas through different cosmological epochs, providing some insight on the corresponding behavior of the HI 21 cm line. We next address the problematic of the (early) enrichment of metals in the IGM, paying attention to the consequences and implications that the chemical enrichment has for models of galaxy formation

and evolution. We also explore the possibility of combining different probes of reionization (like the European Extremely Large Telescope, or the James Web Space Telescope) to gain some more insight into the last episodes of reionization. Finally, at the end of this chapter, we enumerate a number of other cosmological problems that can be addressed (at least partially) with the future SKA radio data.

2 Dark Ages and the End of Reionization

2.1 The HI 21 cm brightness temperature fluctuations

The differential brightness temperature of an hydrogen cloud when observed against a background source of light is [1]:

$$\delta T_b \approx 27\, x_{HI} \left(\frac{T_S - T_R}{T_S}\right) (1+\delta_b) \left(\frac{\Omega_b h^2}{0.023}\right) \left(\frac{0.15}{\Omega_m h^2}\frac{1+z}{10}\right)^{1/2} \left[\frac{\partial_r v_r}{(1+z)H\,(z)}\right] \mathrm{mK}, \qquad (1)$$

where x_{HI} is the fraction of neutral hydrogen, δ_b is the fractional overdensity in baryons and the final term arises from the velocity gradient along the line of sight $\partial_r v_r$. The temperatures T_S and T_R are the *spin temperature* of the gas and the *brightness temperature* of the background radiation, respectively. Thus, the differential brightness temperature of the 21 cm is very sensitive to the environment and physical state of the intergalactic medium (IGM), as well as to fundamental cosmology, that enters in the last four terms of equation (1).

The interplay between temperatures T_S and T_R in equation (1) determines whether the differential temperature brightness δT_b is seen as an absorption or an emission against the backlight CMB radiation. The spin temperature interpolates between the values of the radiation bath T_S and the kinetic temperature T_K of the IGM; therefore, there will be a non-null signal only when T_S couples with T_K. This coupling can appear due to two physical processes: collisions, which are effective in the IGM at high redshifts, $z \geq 50$, or resonant coupling with a Lyman alpha background (Wouthuysen-Field coupling [2]), effective soon after the first sources (stars and galaxies) turn on at lower redshifts. Other second-order effects that may affect T_S include heating by the decay of exotic dark matter particles or photon cascades originated by background X-rays. Although the exact timing of the cosmic epochs is uncertain, the relative order is robustly predicted [3]: the signal transitions from an early phase of collisional coupling to a later phase of Lyα coupling through a short period where there is little signal. Fluctuations after this phase are dominated successively by spatial variation in the Lyman α, X-ray, and ionizing UV radiation backgrounds. After reionization is complete there is a residual signal from neutral hydrogen in galaxies. The net effect is a decrement of background brightness temperature during the dark ages and an increment during the EoR. Somewhere during the twilight of the first galaxies, the 21 cm signal virtually disappears. Though this qualitative behaviour is well understood, the quantitative details of δT_b and the physics behind it are almost unknown. Figure 1 shows this qualitative behaviour as a function of time, with specification of the main milestones of cosmic history.

The Cosmic Dawn and the EoR are roughly marked by the redshift z_* at which the first stars start forming, the redshift z_α at which the Lyman α coupling saturates, the redshift

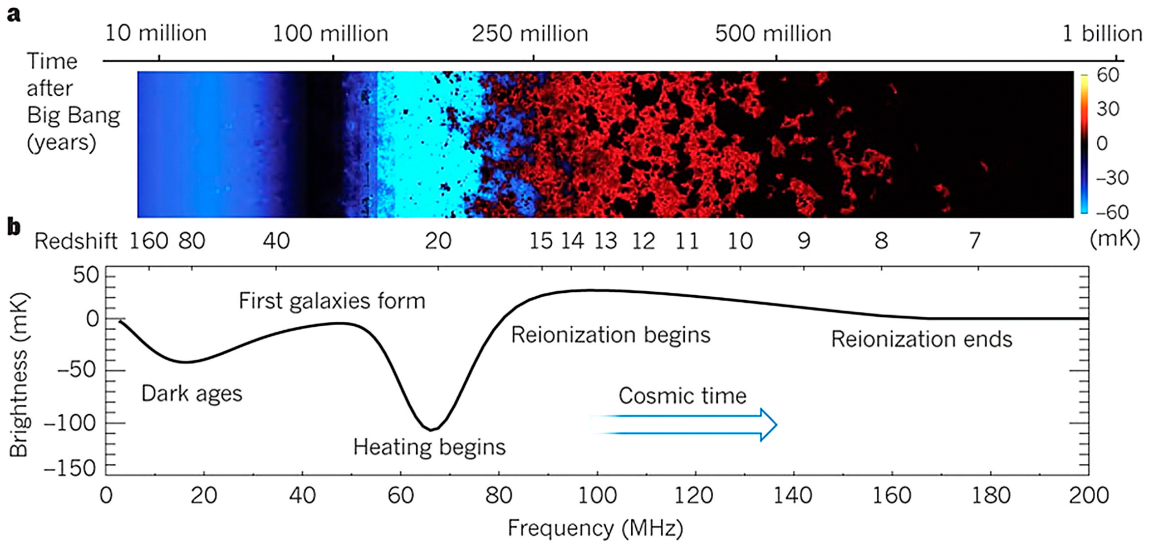

Figure 1: The 21 cm cosmic hydrogen signal. (a) Time evolution of fluctuations in the 21 cm brightness from just before the first stars formed through to the end of the reionization epoch. Coloration indicates the strength of the 21 cm brightness as it evolves through two absorption phases (purple and blue), separated by a period (black) where the excitation temperature of the 21 cm hydrogen transition decouples from the temperature of the hydrogen gas, before it transitions to emission (red) and finally disappears (black) owing to the ionization of the hydrogen gas. (b) Expected evolution of the sky–averaged 21 cm brightness from the Dark Ages at redshift 200 to the end of reionization, sometime before redshift 6. There is considerable uncertainty in the exact form of this signal, arising from the unknown properties of the first galaxies. Reproduced with permission from [4].

z_h at which the gas temperature is heated to the temperature of the CMB, the redshift z_T at which the 21 cm brightness temperature saturates, and the redshift z_r at which the IGM is fully reionized. Most of these milestones are not sharply defined, and so there could be considerable overlap between them. In fact, our ignorance of early sources is such that we cannot definitively be sure of the sequence of events. Following [1] and references therein, we can summarize the approximate sequence of events and observational effects as:

$z_\alpha \leq z \leq z_*$: the spin temperature is still coupled to cold gas and it is lower than T_R and therefore there is an absorption signal. Fluctuations are dominated by density fluctuations and variation in the Lyα flux.

$z_h \leq z \leq z_\alpha$: fluctuations in Lyα flux no longer affect the 21 cm signal and the spin temperature is driven by the kinetic temperature of the atoms. The signal is still in absorption, but as the temperature increases hotter regions may begin to be seen as islands in emission.

$z_T \leq z \leq z_h$: after a transition of approximately no 21 cm signal anywhere, the spin temperature surpasses the value of T_R and we expect to see a 21 cm in emission. Brightness temperature fluctuations are sourced by a mixture of fluctuations in ionization, density and gas temperature. By the end of this phase, the ionization fraction has likely risen above the per cent level.

$z_r \leq z \leq z_T$: temperature fluctuations become unimportant. By this point, the filling fraction of HII regions probably becomes significant and ionization fluctuations begin to dominate the 21 cm signal.

$z \leq z_r$: after reionization, any remaining 21 cm signal originates primarily from collapsed islands of neutral hydrogen (damped Lyα systems).

As noted by [5], the X-ray background due to the appearance of the first stars may give origin to a significant number of Lyα photons that could complicate the previous scenarios even more. Additionally, shocks could raise the IGM temperature very early and exotic particle physics mechanisms such as dark matter annihilation can also play a role. Clearly, there is still considerable uncertainty in the exact evolution of the signal making the potential implications of measuring the 21 cm signal very exciting.

2.2 The history of metal enrichment

It is during the epoch of reionization when metals are produced and ejected to the IGM. How this proceeds is intimately related to the process of galaxy formation and evolution, so whatever description is adopted must be consistent with the very different observational constraints existing not only on metallicities at high redshifts in different environments, but also in other properties of galaxies such as the fraction of visible matter or the metallicity inside galaxies.

By taking spectra along the direction of high redshift quasars, it is now well established that the metallicity of the IGM depends on the local gas density (and it tends to correlate

with it, particularly in the denser environments, (see [6] and citations to this work). There is also evidence that metallicities remain flat versus redshift in the interval $z \in [2, 5]$, decreasing with increasing/decreasing redshifts outside this range, [6, 7, 8]. A satisfactory explanation for these observations must also explain why the fraction of visible matter in galaxies tends to decrease with increasing stellar halo mass, and why the metallicity in halos increases with stellar mass up to a *plateau* found at $M_\star \sim 10^{10}\,\mathrm{M}_\odot$.

A popular explanation nowadays invokes the *pre-enrichment* scenario, by which most of the reionization is due to Population III stars (with no metallicity, formed out of pristine gas) placed in small-mass galaxies ($M_\star < 10^7\,\mathrm{M}_\odot$) that were able to finish reionization by $z \sim 7$. Such galaxies would lose their metals early on due to supernova feedback, thus increasing the metallicity of the IGM while more massive galaxies above a given threshold, with a deeper gravitational potential well, would retain practically all their metals (in agreement with observations). As a consequence of this *feedback*, small halos would have a lower fraction of gas and visible matter too.

SKA observations will not provide measurements of the metallicities in the IGM and hence will not be able to confront this scenario directly. HI 21 cm observations will rather provide a tomographic view of the HI distribution, which constitutes *complementary* information. However, it is thought that by conducting intensity mapping (hereafter IM) on the HI 21 cm line it will be possible to identify those shielded, high density HI reservoirs placed at the cores of the ionized regions and hosting the ionizing sources. The definitive piece of information about these cores and the metal distribution may come from the synergy with other experiments probing reionization, either in the optical or infrared, or even in the millimeter (deferred to the chapter of this book by Diego et al. (Synergies with the Cosmic Microwave Background radiation)).

2.3 Synergies with optical and infrared probes of reionization

As of today there is a number of projects involving the construction of infrared and optical telescopes (some of them of gigantic size) such as the (optical) European Extremely Large Telescope (E-ELT, [9]), or the (infrared) James Webb Space Telescope (JWST, [10]). These facilities have been designed to detect bright galaxies placed as far as $z \in [10, 15]$ and so sample the last moments of the epoch of reionization. However, even when instruments of this size will only be able to detect a small fraction (the brightest tail) of the galaxies present at those redshifts, they should be able to identify and resolve the knots or the centres of the ionized regions sampled at the end of the EoR by HI 21 cm observations. This would provide a complementary description of the universe at those redshifts, since the smooth, large scale neutral and ionized regions would already be mapped by the HI 21 cm data.

Also complementary is the ability of the Atacama Large Millimeter/
Submillimeter Array (hereafter ALMA, [11]) to map molecular emission from clouds hosting young stars in the last stages of reionization ($z \in [8, 10]$). It is possible to combine the spectral and spatial information in SKA and ALMA data in order capture different lines observed at different frequencies but corresponding at the same redshift and to the same pixel on the sky (i.e., the same galaxy). Some of these molecular lines, such as those due to CO rotational

levels, would set further constraints on the history of metal enrichment of the IGM.

3 Before the Cosmic Dawn: recombination lines

Another important predicted signal in the all-sky mean spectrum of the cosmic radio background is the presence of recombination lines from the epoch of cosmological recombination [27, 28] at redshifts $500 \leq z \leq 2000$ for HI, $1600 \leq z \leq 3500$ for HeII $\rightarrow$ HeI, and $5000 < z < 8000$ for HeIII $\rightarrow$ HeII recombination (see e.g. [29] and references therein). In contrast to the 21 cm signatures of reionization, the physical ingredients to describe the epoch of cosmological recombination are simple and well-understood. This fact allows us to find, within the standard model, potentially measurable consequences for the observed energy spectrum [30, 31, 32], as shown in Figure 2. These features could be used: (a) to measure in a model-independent way the redshift of the cosmological recombination; (b) to derive cosmological parameters (e.g. baryon fraction, Hubble parameter, etc.) in a new and completely independent way; and (c) to determine the pre-stellar abundance of helium in the Universe.

As shown in Figure 2, recombinational lines are expected to cover all SKA bands. However, the optimum frequency range for a SKA detection is at the top of the SKA-mid band, where the Galactic contamination is lower and the relative contrast of the distortions with respect to the un-distorted black-body spectrum is still of the order of 10^{-7}. Being an all-sky (monopole) signal, this detection could be achieved with the total power spectra of the dish elements that form the SKA-mid interferometer [34]. Another possibility to measure all-sky signals interferometrically uses the presence of an occulting object (such as the Moon) that imposes spatial structure on an otherwise featureless all-sky signal [33, 34].

4 Other science cases: new sources of energy injection in the Universe

There are multiple physical processes that could inject energy into the IGM, whose effect could be potentially detectable via observations of the distortions of the CMB black body spectrum and the 21 cm line. In this section we focus on those detectable in the radio range. Among others, we can mention here:

- Silk damping of small-scale perturbations gives rise to CMB spectral distortions [12, 13, 14], providing information on the shape and amplitude of the primordial power spectrum at scales $0.6\,\mathrm{kpc} < \lambda < 1\,\mathrm{Mpc}$ (or CMB multipoles between 10^5 and 10^8), thus allowing us to probe 10 additional e-folds of inflation with respect to CMB anisotropies. Measuring this signature also provides an independent approach to study the scale-dependence of f_{NL}, the non-Gaussian nature of the primordial fluctuations [15]. These features are again monopole signals, potentially observable in the SKA-mid band.
- Decay or annihilation of relic particles. The CMB spectrum can be used to set tight limits on decaying and annihilating particles during the pre-recombination epoch [16,

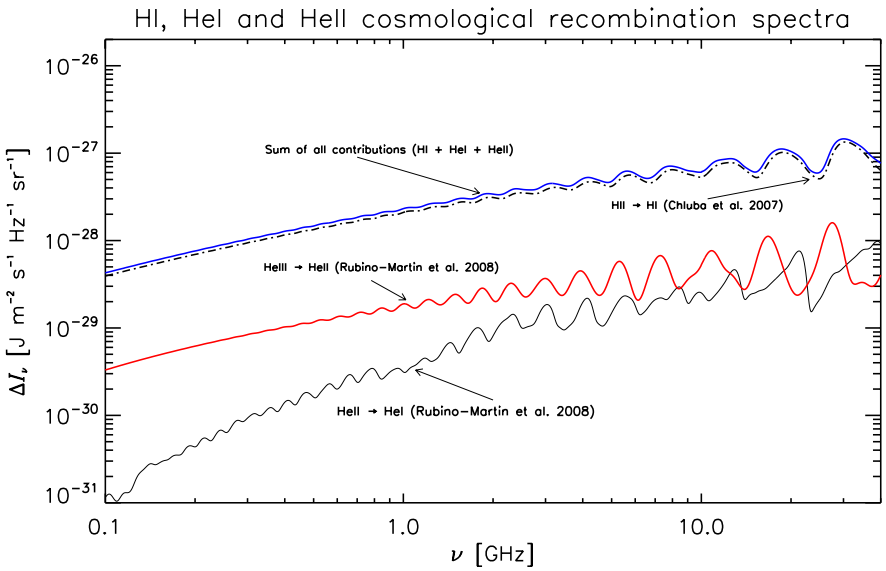

Figure 2: Full helium and hydrogen (bound-bound) cosmological recombination spectra, in the observing window of SKA. Details on the physical modelling of each spectrum can be found in [31, 32] and references therein.

17]. The constraints are especially interesting for decaying particles with lifetimes $\sim 10^8$–10^9 s [18, 19], providing a unique test of the early-universe particle physics. Note that the CMB spectrum is less sensitive at particle masses higher than $m_{\rm DM} \sim 50\,{\rm GeV}$ and for decaying DM [20, 21], while the 21 cm line is more appropriate to explore this regime due to its sensitivity to smaller (and later time) energy injections [22].

- Evaporating black holes. Several inflationary scenarios predict the existence of primordial black holes (PBH) via originated after the collapse of overdense peaks in the initial density field. Those PBHs with masses in the range 5×10^{13}–10^{17} g will evaporate due to Hawking radiation, and thus they will inject energy in the IGM. This excess of energy could be detectable as additional spin temperature fluctuations [23]. Note that for the larger masses, the expected signal will mimic that of a decaying dark matter species.

- Constraints on the radio-dim extragalactic population. There is currently some controversy on the interpretation of the results of the experiment ARCADE 2 [24, 25], since there is apparently an excess of radio background that cannot be explained with the CMB. This excess could be due to a high-redshift, faint radio source population placed at redshifts $z > 1$, but could also be due to other non-thermal mechanism of energy injection into the IGM. The SKA will be able to produce estimates of the radio source luminosity function down to very dim fluxes (5 to 200 nJy), and hence it should be able to confirm or discard the possibility of a discrete radio source population being behind the ARCADE 2 excess. At the same time, it should also set constraints on the free-free emission in halos throughout cosmic history [26], and improve current constraints on

free-free induced distortion [25].

Acknowledgments

The authors acknowledge partial support from the Spanish *Ministerio de Economía y Competitividad* (MICINN) through projects AYA2013-48623-C2-2, AYA2007-68058-C03-01, AYA2010-21766-C03-02, AYA2012-30789, and the Consolider-Ingenio project CSD2010-00064 (EPI: Exploring the Physics of Inflation). We also acknowledge the support of the *Ramón y Cajal* fellowship (RyC 2011 148062) awarded by the Spanish MICINN and the *Marie Curie Career Integration Grant* (CIG 294183).

References

[1] Pritchard, J. and Loeb, A., 2012, Rep. Prog. Phys., 75, 086901

[2] Wouthuysen, S. A., 1952, The Astronomical Journal 57, 31

[3] Furlanetto, S. R., 2006, MNRAS 371, 867

[4] Pritchard, J. and Loeb, A., 2010, Nature 468, 772

[5] Chen, X. & Miralda-Escudé, J. 2008, ApJ, 684, 18

[6] Schaye, J., Aguirre, A., Kim, T.-S. et al. 2003, ApJ, 596, 768

[7] D'Odorico, V., Calura, F., Cristiani, S., & Viel, M. 2010, MNRAS, 401, 2715

[8] D'Odorico, V., Cupani, G., Cristiani, S., et al. 2013, MNRAS, 435, 1198

[9] Shearer, A., Kanbach, G., Slowikowska, A., et al. 2010, Proceedings of High Time Resolution Astrophysics - The Era of Extremely Large Telescopes (HTRA-IV). May 5 - 7, 2010. Agios Nikolaos, Crete, Greece

[10] Greenhouse, M., & JWST Science Working Group 2011, 2010 NASA Laboratory Astrophysics Workshop, 8

[11] Blain, A. W. 2013, New Trends in Radio Astronomy in the ALMA Era: The 30th Anniversary of Nobeyama Radio Observatory, 476, 9

[12] Barrow, J. D., & Coles, P. 1991, MNRAS, 248, 52

[13] Daly, R. A. 1991, ApJ, 371, 14

[14] Hu, W., Scott, D., & Silk, J. 1994, ApJ Lett, 430, L5

[15] Biagetti, M., Perrier, H., Riotto, A., & Desjacques, V. 2013, Phys. Rev. D, 87, 063521

[16] Hu, W., & Silk, J. 1993, Physical Review Letters, 70, 2661

[17] Chluba, J. 2010, MNRAS, 402, 1195

[18] Chluba, J. 2013, MNRAS, 436, 2232

[19] Chluba, J., & Jeong, D. 2014, MNRAS, 438, 2065

[20] Galli, S., Iocco, F., Bertone, G., & Melchiorri, A. 2009, Phys. Rev. D, 80, 023505

[21] Planck Collaboration 2014, A&A, 571, AA16

[22] Furlanetto, S. R., Oh, S. P., & Pierpaoli, E. 2006, Phys. Rev. D, 74, 103502

[23] Mack, K. J. & Wesley, D. H. 2008, eprint arXiv:0805.1531

[24] Singal, J., Fixsen, D. J., Kogut, A., et al. 2011, ApJ, 730, 138

[25] Seiffert, M., Fixsen, D. J., Kogut, A., et al. 2011, ApJ, 734, 6

[26] Ponente, P. P., Diego, J. M., Sheth, R. K., et al. 2011, MNRAS, 410, 2353

[27] Zel'dovich, Y. B., Kurt, V. G., & Sunyaev, R. A. 1968, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 55, 278

[28] Peebles, P. J. E. 1968, ApJ, 153, 1

[29] Sunyaev, R. A., & Chluba, J. 2009, Astronomische Nachrichten, 330, 657

[30] Rubiño-Martín, J. A., Chluba, J., & Sunyaev, R. A. 2006, MNRAS, 371, 1939

[31] Chluba, J., Rubiño-Martín, J. A., & Sunyaev, R. A. 2007, MNRAS, 374, 1310

[32] Rubiño-Martín, J. A., Chluba, J., & Sunyaev, R. A. 2008, A&A, 485, 377

[33] Shaver, P. A., Windhorst, R. A., Madau, P., & de Bruyn, A. G. 1999, A&A, 345, 380

[34] Subrahmanyan, R., Udaya Shankar, N., Pritchard, J., & Vedantham, H. K. 2015, eprint arXiv:1501.04340