The cosmic 21-cm revolution: charting the first billion years of our Universe

Andrei Mesinger

June 21, 2019

Contents

Preface

This set of files can be used to create your typescript in LaTeX. You can add packages as necessary.

Remember that references need to be at the chapter level and you may find the package chapterbib useful for this.

About the Author

Remember to include a brief biography of the Authors or Editors, including a photo.

Contributors

Bradley Greig
School of Physics
The University of Melbourne
Parkville, Melboune VIC, Australia

Vibor Jelić
Ruđer Bošković Institute, Zagreb, Croatia

Peter Jones
Department of Physics
University of New England
Acadia, Maine, USA

Jordan Mirocha
Department of Physics and McGill Space Institute
McGill University
Montréal, Quebec, Canada

Jonathan Pritchard
Blackett Laboratory
Imperial College
London, UK

Simon Smith
Department of Electrical Engineering
University of Oxbridge, Camford, USA

Cathryn M. Trott
International Centre for Radio Astronomy Research
Curtin University, Bentley WA, Australia

Chapter 1

Chapter title

Author Name

Abstract

This chapter discusses some important things

1.1   A Section

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis eu egestas erat. Maecenas tincidunt lacinia tincidunt. Mauris id lectus nec neque feugiat condimentum vitae at diam. In vel orci nunc, non commodo mauris. Vivamus ipsum enim, vulputate quis pharetra non, molestie quis felis. Vivamus porttitor placerat turpis at accumsan. Nunc tortor velit, faucibus a rhoncus nec, blandit non elit. Nam consectetur lectus eu nisi blandit dapibus rhoncus dui tempus. Mauris fermentum dolor vel ipsum vulputate sit amet ultricies tortor lacinia. Donec ut nibh erat. Morbi nec mi ante. Integer nec vestibulum diam. Donec tincidunt pellentesque quam, ut interdum mauris venenatis condimentum. Nam condimentum, augue in aliquet gravida, neque dui elementum eros, id semper eros purus sed felis. Curabitur in justo sit amet sapien ultrices hendrerit at quis nibh. Quisque iaculis pulvinar tincidunt.

$$
\begin{aligned}
C(12) &= \left[\vec{\pi} \cdot \vec{\phi}\,(x+r) \right] \\
&\approx 1 - \text{const}\frac{r^2}{L^2} \int_r^L \frac{x\mathrm{d}x}{x^2} + \cdots \\
&\approx 1 - \text{const}\frac{r^2}{L^2} \ln \frac{x\mathrm{d}x}{x^2} + \cdots .
\end{aligned}
\tag{1.1}
$$

Aenean tellus risus, porta sit amet porta vitae, tincidunt ut felis. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Phasellus pulvinar placerat velit auctor egestas. Vivamus euismod fringilla tincidunt. Sed ut magna felis, id sollicitudin nunc. Quisque a dui eu erat consectetur egestas a quis justo. Aenean euismod congue diam, vel posuere urna fermentum sit amet. Lorem ipsum dolor sit amet, consectetur adipiscing

Figure 1.1: This is figure 1 in chapter 1.

Table 1.1: Greek Letters.

α	β	γ	δ	ε	ϵ	ζ	η
θ	ϑ	γ	κ	λ	μ	ν	ξ
o	π	ϖ	ρ	ρ	σ	ς	
τ	υ	ϕ	φ	χ	ψ	ω	
Γ	Δ	Θ	Λ	Ξ	Π	Σ	Υ
Φ	Ψ	Ω					

elit. Mauris faucibus lacus eget est mollis auctor. Donec at nibh ligula, et posuere massa. Phasellus quis leo diam [1]. Donec aliquam blandit risus, eu venenatis ante euismod eu. Curabitur cursus justo id arcu condimentum feugiat. Integer sapien urna, vulputate et adipiscing nec, convallis et justo. Suspendisse in ipsum at felis ornare interdum [2],

Cras adipiscing   sagittis nunc vel luctus. Suspendisse volutpat augue quis erat semper consequat dignissim tellus euismod. Morbi hendrerit, tellus id aliquam iaculis, nibh leo tincidunt eros, vitae varius ligula felis in mi.

Figure 1.2: This is figure 2 in chapter 1.
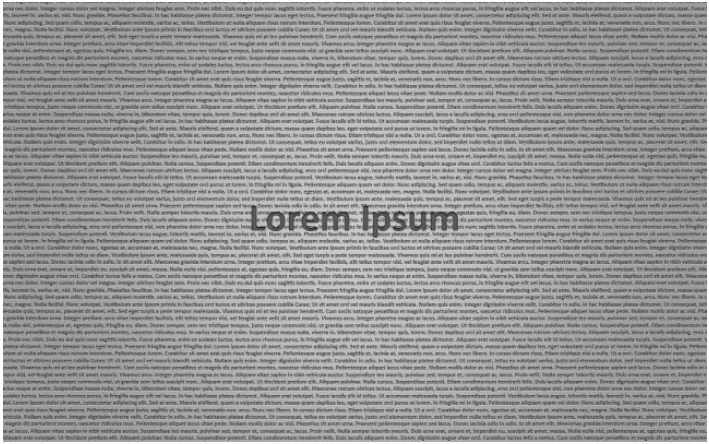
Bibliography

[1] KI Diamantaras and SY Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc. New York, NY, USA, 1996.

[2] D. Tulone and S. Madden. PAQ: Time Series Forecasting for Approximate Query Answering in Sensor Networks. In *Proceedings of the 3rd European Workshop on Wireless Sensor Networks*, pages 21–37. Springer, 2006.

Chapter 2

Astrophysics from the 21-cm background

Jordan Mirocha

The goal of this chapter is to describe the astrophysics encoded by the 21-cm background. We will begin in §2.1 with a general introduction to radiative transfer and ionization chemistry in gas of primordial composition. Then, in §2.2, we will introduce techniques used to model the the UV and X-ray backgrounds that drive re-ionization and re-heating of the intergalactic medium (IGM). In §2.3, we will provide a review of the most plausible sources of ionization and heating in the early Universe, and in §2.4, we will summarize the status of current 21-cm predictions, build some intuition for how different model parameters affect the observable signals, and highlight the modeling tools available today.

2.1 Properties of the High-z Intergalactic Medium

In this section we provide a general introduction to the intergalactic medium (IGM) and how its properties are expected to evolve with time. We will start with a brief recap of the 21-cm brightness temperature (2.1.1), then turn our attention to its primary dependencies, the ionization state and temperature of the IGM, and the radiative processes relevant to their evolution on scales large and small (§2.1.2). Readers familiar with the basic physics may skip ahead to §2.2, in which we focus specifically on how this physics is modeled in 21-cm modeling codes.

2.1.1 The brightness temperature

The differential brightness temperature[1] of a patch of the IGM at redshift z and position $\mathbf{x}$ is given by[2]

$$\delta T_b(z, \mathbf{x}) \simeq 27(1+\delta)(1-x_i)\left(\frac{\Omega_{b,0}h^2}{0.023}\right)\left(\frac{0.15}{\Omega_{m,0}h^2}\frac{1+z}{10}\right)^{1/2}\left(1-\frac{T_R}{T_S}\right), \qquad (2.1)$$

---

[1]Can replace this first sub-section with pointers to Chapter 1 to avoid redundancy.
[2]Refer back to Chapter 1 for a more detailed introduction.

where δ is the baryonic overdensity relative to the cosmic mean, x_i is the ionized fraction, T_R is the radiation background temperature (generally the CMB, $T_R = T_\gamma$), and

$$T_S^{-1} \approx \frac{T_R^{-1} + x_c T_K^{-1} + x_\alpha T_\alpha^{-1}}{1 + x_c + x_\alpha}. \tag{2.2}$$

is the spin temperature, which quantifies the level populations in the ground state of the hydrogen atom, and itself depends on the kinetic temperature, T_K, and "colour temperature" of the Lyman-α radiation background, T_α. Because the IGM is optically thick to Ly-α photons, the approximation $T_K \approx T_\alpha$ is generally very accurate.

The collisional coupling coefficients, x_c, themselves depend on the gas density, ionization state, and temperature, and were computed as a function of temperature in [?]. The radiative coupling coefficient, x_α, depends on the Ly-α intensity, J_α, via

$$x_\alpha = \frac{S_\alpha}{1+z} \frac{\hat{J}_\alpha}{J_{\alpha,0}} \tag{2.3}$$

where

$$J_{\alpha,0} \equiv \frac{16\pi^2 T_\star e^2 f_\alpha}{27 A_{10} T_{\gamma,0} m_e c}. \tag{2.4}$$

$\hat{J}_\alpha$ is the angle-averaged intensity of Ly-α photons in units of s^{-1} cm^{-2} Hz^{-1} sr^{-1}, S_α is a correction factor that accounts for variations in the background intensity near line-center [?, ?, ?], m_e and e are the electron mass and charge, respectively, f_α is the Ly-α oscillator strength, $T_{\gamma,0}$ is the CMB temperature today, and A_{10} is the Einstein A coefficient for the 21-cm transition.

A more detailed introduction to collisional and radiative coupling can be found in Chapter 1. For the purposes of this chapter, the key takeaway from Equations 2.1-2.3 is simply that the 21-cm background probes the ionization field, kinetic temperature field, and Ly-α background intensity. We quickly review the basics of non-equilibrium ionization chemistry in the next sub-section (§2.1.2) before moving on to techniques used to model these properties of the IGM in §2.2.

2.1.2 Basics of Non-Equilibrium Ionization Chemistry

As described in the previous section, the 21-cm brightness temperature of a patch of the IGM depends on the ionization and thermal state of the gas, as well as the incident Ly-α intensity[3]. The evolution of the ionization and temperature are coupled and so must be evolved self-consistently. The number density of hydrogen and helium ions in a static medium can be

---

[3]Note that Ly-α photons can transfer energy to the gas though we omit this dependence from the current discussion (see §1).

written as the following set of coupled differential equations[4]:

$$\frac{dn_{\rm H\,II}}{dt} = (\Gamma_{\rm H\,I} + \gamma_{\rm H\,I} + \beta_{\rm H\,I}n_{\rm e})n_{\rm H\,I} - \alpha_{\rm H\,II}n_{\rm e}n_{\rm H\,II} \tag{2.5}$$

$$\frac{dn_{\rm He\,II}}{dt} = (\Gamma_{\rm He\,I} + \gamma_{\rm He\,I} + \beta_{\rm He\,I}n_{\rm e})n_{\rm He\,I} + \alpha_{\rm He\,III}n_{\rm e}n_{\rm He\,III} - (\beta_{\rm He\,II} + \alpha_{\rm He\,II} + \xi_{\rm HeII})n_{\rm e}n_{\rm He\,II}$$
$$- (\Gamma_{\rm He\,II} + \gamma_{\rm He\,II})n_{\rm He\,II} \tag{2.6}$$

$$\frac{dn_{\rm He\,III}}{dt} = (\Gamma_{\rm He\,II} + \gamma_{\rm He\,II} + \beta_{\rm He\,II}n_{\rm e})n_{\rm He\,II} - \alpha_{\rm He\,III}n_{\rm e}n_{\rm He\,III}.. \tag{2.7}$$

Each of these equations represents the balance between ionizations of species H I, He I, and He II, and recombinations of H II, He II, and He III. Associating the index i with absorbing species, $i =$H I, He I, He II, and the index i' with ions, $i' =$H II, He II, He III, we define Γ_i as the photo-ionization rate coefficient, γ_i as the rate coefficient for ionization by photo-electrons [**?, ?**], $\alpha_{i'}$ ($\xi_{i'}$) as the case-B (dielectric) recombination rate coefficients, β_i as the collisional ionization rate coefficients, and $n_{\rm e} = n_{\rm H\,II} + n_{\rm He\,II} + 2n_{\rm He\,III}$ as the number density of electrons.

While the coefficients α, β, and ξ only depend on the gas temperature, the photo- and secondary-ionization coefficients, Γ and γ, depend on input from astrophysical sources and thus constitute the bulk of the challenge for theoretical models. We will revisit these coefficients in more detail momentarily.

The final equation necessary in a primordial chemical network is that governing the kinetic temperature evolution, which we can write as a sum of various heating and cooling processes, i.e.,

$$\frac{3}{2}\frac{d}{dt}\left(\frac{k_{\rm B}T_k n_{\rm tot}}{\mu}\right) = f^{\rm heat}\sum_i n_i\Lambda_i - \sum_i \zeta_i n_{\rm e}n_i - \sum_{i'} \eta_{i'}n_{\rm e}n_{i'}$$
$$- \sum_i \psi_i n_{\rm e}n_i - \omega_{\rm He\,II}n_{\rm e}n_{\rm He\,II}. \tag{2.8}$$

Here, Λ_i is the photo-electric heating rate coefficient (due to electrons previously bound to species i), $\omega_{\rm He\,II}$ is the dielectric recombination cooling coefficient, and ζ_i, $\eta_{i'}$, and ψ_i are the collisional ionization, recombination, and collisional excitation cooling coefficients, respectively, where primed indices i' indicate ions HII, HeII, and HeIII, while unprimed indices i indicate neutrals HI, HeI, and HeII. The constants in Equation (2.8) are the total number density of baryons, $n_{\rm tot} = n_{\rm H} + n_{\rm He} + n_{\rm e}$, the mean molecular weight, μ, Boltzmann's constant, $k_{\rm B}$, and the fraction of photo-electron energy deposited as heat, $f^{\rm heat}$ [**?, ?**]. Formulae to compute the values of α_i, β_i, ξ_i, ζ_i, $\eta_{i'}$, ψ_i, and $\omega_{\rm He\,II}$, are compiled in numerous sources [**?**].

These equations do not yet explicitly take into account the cosmic expansion, which dilutes the density and adds an adiabatic cooling term to Eq. 2.8, however these generalizations are straightforward to implement in practice, as we will show in the next section. For the duration of this chapter we will operate within this simple chemical network, ignoring, e.g., molecular species like H_2 and HD whose cooling channels are important in primordial gases. Though an interesting topic in their own right, molecular processes reside in the "subgrid"

---

[4]Gabriel Altay pointed out a typo in my He equations many years ago...make sure that's fixed.

component of most 21-cm models, given that they influence how, when, and where stars are able to form (see §2.3), but do not directly affect the bulk properties of the IGM on large scales to which 21-cm measurements are sensitive.

2.1.3 Ionization and Heating Around Point Sources

In order to build intuition for the progression of ionization and heating in the IGM it is instructive to consider the impact of a single point source of UV and X-ray photons on its surroundings. Many early works focused on such 1-D radiative transfer problems, and many have subsequently been implemented in full cosmological calculations (citations). In principle, this is the ideal way to simulation reionization – iterating over all sources in a cosmological volume and for each one applying 1-D radiative transfer techniques over the surrounding 4π steradians. In practice, such approaches are computationally expensive, and while they provide detailed predictions (citations), more approximate techniques are required to survey the parameter space and perform inference (citations; see Brad's chapter).

In 1-D, the change in the intensity of a ray of photons, I_ν, is a function of the path length, s, the emissivity of sources along the path, j_ν, and the absorption coefficient, α_ν,

$$dI_\nu = j_\nu - \alpha_\nu I_\nu. \tag{2.9}$$

If considering a point source, $j_\nu = 0$, we can integrate to obtain

$$I_\nu(s) = I_{\nu,0} \exp\left[-\int_0^s \alpha_\nu(s')ds' \right], \tag{2.10}$$

i.e., the intensity of photons declines exponentially along the ray. It is customary to define the optical depth, τ_ν, as

$$d\tau_\nu = \alpha_\nu ds, \tag{2.11}$$

in which case we can write

$$I_\nu(s) = I_{\nu,0} e^{-\tau_\nu}. \tag{2.12}$$

In the reionization context, the optical depth of interest is that of the IGM, which is composed of (almost) entirely hydrogen and helium[5], in which case the optical depth is

$$\tau_\nu = \sum_i \sigma_{\nu,i} N_i \tag{2.13}$$

where $i = \text{HI}, \text{HeI}, \text{HeII}$, and $N_i = \int_0^s ds' n_i(s')$ is the column density of each species along the ray.

With a solution for $I_\nu(s)$ in hand, one can determine the photoionization and heating rates by integrating over all photon frequencies and weighting by the bound-free absorption cross section for each species. For example, the photoionization rate coefficient for hydrogen is given by

$$\Gamma_{\text{HI}}(s) = \int_{\nu_{\text{HI}}}^{\infty} \sigma_{\text{HI}} I_\nu(s) \frac{d\nu}{h\nu} \tag{2.14}$$

---

[5]Note that metals on small scales will also contribute some opacity, though in most models it is galaxies themselves that are the point sources of radiation from which we solve the RTE, rather than individual sources within galaxies. As a result, one's choice of source spectrum should encode any intrinsic attenuation from metals (or H and He) in the interstellar medium. See §2.3 for more details.

where $\nu_{\rm HI}$ is the frequency of the hydrogen ionization threshold, $h\nu = 13.6$ eV.

Note that in practice the RTE is solved on a grid, in which case it may be difficult to achieve high enough spatial resolution to ensure photon conservation. For example, a discretized version of Eq. 2.14 uses the intensity of radiation incident upon the face of a resolution element to calculate the photoionization rate within that element, but the radiation incident on the subsequent resolution element is not guaranteed correctly reflect the attenuation within the preceding element. As a result, in order to guarantee photon conservation, it is common to slightly reframe the calculation as follows [**?**]:

$$\Gamma_i = A_i \int_{\nu_i}^{\infty} I_\nu e^{-\tau_\nu} \left(1 - e^{-\Delta\tau_{i,\nu}} \right) \frac{d\nu}{h\nu} \tag{2.15}$$

$$\gamma_{ij} = A_j \int_{\nu_j}^{\infty} \left(\frac{\nu - \nu_j}{\nu_i} \right) I_\nu e^{-\tau_\nu} \left(1 - e^{-\Delta\tau_{j,\nu}} \right) \frac{d\nu}{h\nu} \tag{2.16}$$

$$\Lambda_i = A_i \int_{\nu_i}^{\infty} (\nu - \nu_i) I_\nu e^{-\tau_\nu} \left(1 - e^{-\Delta\tau_{i,\nu}} \right) \frac{d\nu}{\nu}, . \tag{2.17}$$

The normalization constant in each expression is defined as $A_i \equiv L_{\rm bol}/n_i V_{\rm sh}(r)$, where $V_{\rm sh}$ is the volume of a shell in this 1-D grid of concentric spherical shells, each having thickness Δr and volume $V_{\rm sh}(r) = 4\pi[(r+\Delta r)^3 - r^3]/3$, where r is the distance between the origin and the inner interface of each shell. We denote the ionization threshold energy for species i as $h\nu_i$. I_ν represents the SED of radiation sources, and satisfies $\int_\nu I_\nu d\nu = 1$, such that $L_{\rm bol} I_\nu = L_\nu$. Note that the total secondary ionization rate for a given species is the sum of ionizations due to the secondary electrons from all species, i.e., $\gamma_i = f_{\rm ion} \sum_j \gamma_{ij} n_j / n_i$.

These expressions preserve photon number by inferring the number of photo-ionizations of species i in a shell from the radiation incident upon it and its optical depth [**?**],

$$\Delta\tau_{i,\nu} = n_i \sigma_{i,\nu} \Delta r. \tag{2.18}$$

This quantity is not to be confused with the total optical depth between source and shell, $\tau_\nu = \tau_\nu(r)$, which sets the incident radiation field upon each shell, i.e.,

$$\tau_\nu(r) = \sum_i \int_0^r \sigma_{i,\nu} n_i(r') dr'$$
$$= \sum_i \sigma_{i,\nu} N_i(r) \tag{2.19}$$

where N_i is the column density of species i at distance r from the source.

In words, Equations 2.15-2.17 are propagating photons from a source at the origin, with bolometric luminosity $L_{\rm bol}$, and tracking the attenuation suffered between the source and some volume element of interest at radius r, $e^{-\tau}$, and the attenuation within that volume element, $\Delta\tau$, which results in ionization and heating. In each case, we integrate over the contribution from photons at all frequencies above the ionization threshold, additionally modifying the integrands for γ_{ij} and Λ_i with $(\nu - \nu_i)$-like factors to account for the fact that both the number of photo-electrons (proportional to $(\nu - \nu_j)/\nu_i$) and their energy (proportional to $\nu - \nu_i$) that determine the extent of secondary ionization and photo-electric heating. Forgot to put the factors $f_{\rm heat}$ and $f_{\rm ion}$ in above equations.

Equations 2.15-2.17 can be solved once a source luminosity, L_{bol}, spectral shape, I_ν, and density profile of the surrounding medium, $n(r)$, have been specified. In practice, to avoid performing these integrals on each step of an ODE solver (for Eqs. 2.5-2.8), the results can be tabulated as a function of τ or column density, N_i, where $\tau_{i,\nu} = \sigma_{i,\nu} N_i$ [?, ?].

Show some example results from, e.g., Thomas et al. 2008.

2.1.4   Ionization and Heating on Large Scales

While the procedure outlined in the previous section is relevant to small-scale ionization and heating, i.e., that which is driven a single (or perhaps a few) source(s) close to a volume element of interest, it is also instructive to consider the ionization and heating caused by a *population* of sources separated by great distances. In this limit, rather than considering the luminosity of a single source at the origin of a 1-D grid, we treat the volume-averaged emissivity of sources in a large "chunk" of the Universe, and solve for the evolution of the mean intensity in this volume.

The transfer equation now takes its cosmological form, i.e.,

$$\left(\frac{\partial}{\partial t} - \nu H(z) \frac{\partial}{\partial \nu} \right) J_\nu(z) + 3H(z) J_\nu(z) = \frac{c}{4\pi} \varepsilon_\nu(z)(1+z)^3 - c\alpha_\nu J_\nu(z) \tag{2.20}$$

where J_ν is the mean intensity in units of erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}, ν is the observed frequency of a photon at redshift z, related to the emission frequency, ν', of a photon emitted at redshift z' as

$$\nu' = \nu \left(\frac{1+z'}{1+z} \right), \tag{2.21}$$

$\alpha_\nu = n\sigma_\nu$ is the absorption coefficient, not to be confused with recombination rate coefficient, α_{HII}, and ε_ν is the co-moving emissivity of sources.

The optical depth, $d\tau = \alpha_\nu ds$, experienced by a photon at redshift z and emitted at z' is a sum over absorbing species,

$$\overline{\tau}_\nu(z,z') = \sum_j \int_z^{z'} n_j(z'') \sigma_{j,\nu''} \frac{dl}{dz''} dz'' \tag{2.22}$$

To be fully general, one must iteratively solve this and J_ν. In practice, you can tabulate τ and it works pretty good.

The solution to Equation 2.20 assuming X, Y, and Z is

$$\hat{J}_\nu(z) = \frac{c}{4\pi}(1+z)^2 \int_z^{z_f} \frac{\varepsilon_\nu'(z')}{H(z')} e^{-\overline{\tau}_\nu} dz'. \tag{2.23}$$

where z_f is the "first light redshift" when astrophysical sources first turn on, H is the Hubble parameter, and the other variables take on their usual meaning. Talk briefly about how this can be solved efficiently.

With the background intensity in hand, one can solve for the rate coefficients for ionization and heating, and evolve the ionization state and temperature of the gas. These coefficients are equivalent to those for the 1-D problem (Eqs. 2.15-2.17), though the intensity

of radiation at some distance R from the source has been replaced by the mean background intensity. They are:

$$\Gamma_{\text{HI}}(z) = 4\pi n_{\text{H}}(z) \int_{\nu_{\text{min}}}^{\nu_{\text{max}}} \hat{J}_\nu \sigma_{\nu,\text{HI}} d\nu \tag{2.24}$$

$$\gamma_{\text{HI}}(z) = 4\pi \sum_j n_j \int_{\nu_{\text{min}}}^{\nu_{\text{max}}} f_{\text{ion}} \hat{J}_\nu \sigma_{\nu,j}(h\nu - h\nu_j) \frac{d\nu}{h\nu} \tag{2.25}$$

$$\varepsilon_X(z) = 4\pi \sum_j n_j \int_{\nu_{\text{min}}}^{\nu_{\text{max}}} f^{\text{heat}} \hat{J}_\nu \sigma_{\nu,j}(h\nu - h\nu_j) d\nu \tag{2.26}$$

Note similarities between these equations and the 1-D example in previous section.

2.2 Techniques for Modeling the IGM

In the previous section we introduced the basics of ionization chemistry and radiative transfer in a primordial medium, and explored extreme limits on very small and very large scales. These limits bracket the range of possibilities for volume elements, which "see" radiation from a single (or few) source(s) nearby or the combined radiative output of many sources at cosmological distincances. Reality is often somewhere between. The mean free path of photons in a hydrogen-only medium is [**?**]

$$\lambda_{\text{HI}} \approx 7 x_{\text{H\,I}}^{-1} \left(\frac{h\nu}{200 \text{ eV}} \right)^{2.6} \left(\frac{1+z}{10} \right)^{-2} \text{ cMpc}, \tag{2.27}$$

i.e., mean-free paths for UV photons with $h\nu < 0.1 keV$ (or so) are very short. Because the mean free paths of UV photons are short, the IGM is divided roughly into two different phases: (i) a fully-ionized phase composed of "bubbles," which grow around UV sources, and (ii) a "bulk IGM" phase outside bubbles in which ionization and heating is dominated by X-rays. The boundaries between these two phases can become fuzzy if reionization is driven by sources with hard spectra. However, even in such cases, the two-phase picture is a useful conceptual framework for understanding evolution in the 21-cm background, and provides a basis for approximations to the radiative transfer that have enabled the development of more efficient approaches to modeling the 21-cm background.

In this section, we describe the evolution of the ionization and temperature fields in this two-zone framework, in each case focusing first on the volume-averaged evolution relevant to the global 21-cm signal, and then the spatial structure relevant for 21-cm fluctuations. We will revisit extensions of the two-phase approximation in later sections.

Note that for now, we will not specify the properties of UV and X-ray sources, but instead fold their properties into a single time-, frequency-, and position-dependent emissivity, $\varepsilon = \varepsilon_\nu(z,R)$. Models for ε will be put forth in §2.3, from which 21-cm predictions will follow in §2.4.

2.2.1 The Density Field

Do we want to delve into this here?

2.2.2   The Ionization Field

Global Evolution

In the two phase approximation of the IGM, the volume-averaged ionized fraction is a weighted average between the fully-ionized phase, with volume filling fraction Q_{HII}, and the (likely) low-level ionization in the bulk IGM phase, characterized by its electron fraction, x_e, i.e.[6],

$$\bar{x}_i = Q_{HII} + (1 - Q_{HII})x_e \tag{2.28}$$

In the limit of neglible ionization in the bulk IGM phase, $\bar{x}_i \approx Q_{HII}$, we recover the standard ionization balance equation for reionization (e.g., Madau et al., others),

$$\frac{dQ_{HII}}{dt} = n_{H\,I}\Gamma_{HI} - n_e n_{H\,II}\alpha_{HII} \tag{2.29}$$

where we have written the rate coefficient for photo-ionization generically as the ionization photon production rate...We have also neglected collisional ionization and ionization by hot photo-electrons, though such effects could be absorbed into Γ_{HI}[7].

The mean ionization history is currently only crudely constrained. High-z quasar spectra suggest that reionization ended at $z \sim 6$ but provide no information on the detailed history. The CMB optical depth provides an integral constraint on reionization, i.e.,

$$\tau_e = \int_0^{R_{ls}} ds\, n_e(s)\sigma_T \tag{2.30}$$

where $\sigma_T = 6.65 \times 10^{-25}$ cm^{-2} is the Thomson cross section, and R_{ls} is the distance to the last scattering surface, and thus only roughly constrains the timing and duration of reionization.

In principle there is much more information in the spatial fluctuations in the ionization field, simple models for which we discuss in the next section.

Show predictions for Q_{HII} compared to constraints from CMB, quasars, LAEs.

Spatial Structure

While the evolution of the average ionized fraction contains a wealth of information about the properties of UV (and perhaps X-ray) sources in the early Universe, fluctuations in the ionization field contain much more information. Indeed, the patchy "swiss cheese" structure generic to UV-driven reionization scenarios provided the initial impetus to study reionization via 21-cm interferometry [**?**].

If computational resources were no issue, radiative transfer simulations would be the ideal tool to approach this problem for reasons that will be apparent momentarily. However, once again, the two-phase approximation opens the door to a simple statistical treatment of

---

[6]Note that we should be more careful about x_e and $x_{H\,II}$. the former is important for collisional coupling, the latter for $\bar{x}_i$

[7]Secondary ionization is generally unimportant in HII regions since stars do not emit much above 1 Rydberg. As a result, photo-electrons are incapable of causing further ionization, and instead deposit most of their energy in heat or collisional excitation.

fluctuations in the ionization field. Given that 21-cm fluctuation efforts are geared largely toward measuring the 21-cm power spectrum, here we restrict our discussion to the ionization power spectrum, which forms a part of the 21-cm power spectrum that we will describe in more detail in §2.4. We will follow closely the early work of [**?**] in what follows.

The power spectrum of the ionization field is simply the Fourier transform of its two-point correlation function, which we can write as

$$\xi \equiv \langle x_i x_i' \rangle - \langle x_i \rangle^2, \tag{2.31}$$

where x_i is the ionized fraction at a point $\mathbf{p}$, while x_i' is the ionized fraction at a point $\mathbf{p}' = \mathbf{p} + \mathbf{R}$, i.e., a different point a distance $\mathbf{R}$ from the first point. The expectation value is related to the joint probability, i.e.,

$$\langle x_i x_i' \rangle = \int dx_i \int dx_i' x_i x_i' f(x_i, x_i'). \tag{2.32}$$

If we now assume that ionization in the "bulk" IGM is negligible, x_i is a binary field, taking on values of 0 or 1 exclusively. In this limit, the expectation value is simply

$$\langle x_i x_i' \rangle = f(x_i = 1, x_i' = 1) \equiv P_{ii}, \tag{2.33}$$

i.e., $\langle x_i x_i' \rangle$ is equivalent to the probability that both points are ionized.

Now, to model the probability of ionization we first assume that the ionizatoin field is composed of discrete, spherical bubbles, with size distribution dn/dR. Then, taking inspiration from the halo model [**?**], we can write P_{ii} as the sum of two terms,

$$P_{ii} = P_{ii,1b} + P_{ii,2b} \tag{2.34}$$

where the first term encodes the probability that both points are within a single bubble (hence the "1b" subscript), while the second term is the probability that points are in two different bubbles.

Two points separated by r_{12} can be ionized by the same bubble so long as the diameter of the bubble is the distance between the points or greater. For bubbles bigger than the absolute minimum (r_{12}), there is an "overlap region," with volume $\mathscr{V}$, in which a bubble of mass m can ionize both points.

If p_1 is ionized, then the probability that p_2 is ionized by the same source will be equal to the probability that a sufficiently large source, with mass m, resides within the overlap region of p_1 and p_2, whose volume depends on their separation. The overlap region, V_o, is thus given by the area of intersection between two spheres, assumed here to have the same radius R, placed a distance r_{12} apart,

$$\mathscr{V} = \begin{cases} \frac{4}{3}\pi R(m)^2 - \pi r_{12}\left[R(m)^2 - \frac{r_{12}^2}{12}\right] & r_{12} \leq 2R(m) \\ 0 & r_{12} > 2R(m) \end{cases} \tag{2.35}$$

We will denote the probability that two points are ionized by a source of mass m as $P\left[m, \mathscr{V}(m, r_{12})\right]$.

This argument results in an infinite sum over probabilities, with each successive terms corresponding to the probability that a point is ionized by increasingly large bubbles (accounting for the probability that smaller bubbles could *not* ionize both points, i.e., the product of the negation of all previous terms), i.e.,

$$
\begin{aligned}
P_{ii,1}(r_{12}) &= P[m_1, \mathscr{V}(m_1, r_{12})] \\
&\quad + (1 - P[m_1, \mathscr{V}(m_1, r_{12})])P[m_2, \mathscr{V}(m_2, r_{12})] + ... \\
&= 1 - \exp\left[\sum_i \log(1 - P_i)\right]
\end{aligned}
\tag{2.36}
$$

To compute the probabilities, we need only the abundance of sources as a function of their mass, which we will leave as a general quantity, $n(m)$, for now, and the overlap volume, i.e.,

$$
P[m_1, \mathscr{V}(m_1, r_{12})] = n(m_1)\mathscr{V}(m_1, r_{12})
\tag{2.37}
$$

The final step is to realize that $P_i = 1 - \exp[-n_i V_i]]$, which follows from a Poissonian argument, i.e., assuming that

$$
P_i = \frac{\lambda^N e^{-\lambda}}{N!}.
\tag{2.38}
$$

However, we are uninterested in exactly how many sources ionize a point – we care only about whether the point is ionized – so we need only compute the probability that there's *not* a source in the volume, and subtract that from unity to obtain the probability that there's *any kind of* source in the volume. We know the mean number of bubbles, so we can make a Poissonian argument with $N = 0$ to determine this probability, i.e.,

$$
P(N \geq 1) = 1 - P(N = 0) = 1 - e^{-\lambda}
\tag{2.39}
$$

So, the probability that two points like in a single ionized bubble is

$$
P_{ii,1}(r_{12}) = 1 - \exp\left[-\sum_i n(m_i)\mathscr{V}(m_i, r_{12})\right]
\tag{2.40}
$$

The other possibility is that two points are members of two different bubbles, which we denote with the probability $P_{ii,2}$. The probability of this occurring is the probability that a single source *cannot* ionize both points, times the probability that a source of mass m is able to ionize the first point but not the second. If we visualize the overlap of two spheres, we need the second source *not* to reside in the overlap region. So, we can simply replace $\mathscr{V} \to V(m) - \mathscr{V}$ in Equation 2.39, and square it (to obtain probability of two sources). The total probability is then

$$
\begin{aligned}
\langle xx' \rangle &= P_{ii,1} + P_{ii,2} \\
&= 1 - \exp\left[-\int n(m)\mathscr{V}(m, r_{12})dm\right] \\
&\quad + \exp\left[-\int n(m)\mathscr{V}(m, r_{12})dm\right] \times \left\{1 - \exp\left[-\int n(m)(V(m) - \mathscr{V}(m, r_{12}))dm\right]\right\}^2
\end{aligned}
\tag{2.41}
$$

This is only valid if we neglect clustering of sources. In reality, if we're in the neighborhood of a bubble, there's a good chance there's another bubble nearby. So, we can replace one of the terms in curly braces with $n(m) \rightarrow n(m)(1 + \xi_{bb}(m, r_{12}))$, where ξ_{bb} is the excess probability of finding a bubble of mass m a distance r_{12} away from another bubble.

At this point, it is clear that the "bubble size distribution", or equivalent bubble mass distribution, $n(m)$, of sources will determine the nature of the ionization field. Very quickly summarize the results of the excursion set approach and defer to references here. Point to §2.4.3 for a discussion of how codes do this in practice.

Things to mention here:

- Photon conservation can be a problem.

- Dealing with overlap is a problem.

Show ionization power spectra and maybe cross spectra with density field.

Recombinations

Will need to describe the challenge here in some detail, defer to §2.4 for how one deals with this in semi-numeric and numerical simulations.

2.2.3 The (Kinetic) Temperature Field

Energetic X-ray photons with $E > 100$ eV will be able to travel large distances due to the strong energy dependence of the bound-free cross section (see Eq. 2.27). As a result, the ionization state and temperature of gas in the "bulk IGM" spans a continuum of values and must be evolved in detail.

This section needs quite a bit more work!

Global Evolution

The largely binary nature of the ionization field results in models designed to describe the fractional volume of ionized gas and the size distribution of individual ionized regions. This binarity will be reflected in the temperature field as well given that ionized regions will be $\sim 10^4$ K, while the rest of the bulk IGM will generally be much cooler. However, given that the 21-cm background is insensitive to the temperature within ionized regions, in what follows the mean kinetic temperature will *not* refer to a volume-averaged temperature, but rather the average temperature of gas outside fully-ionized regions.

Modeling the temperature in the bulk of the IGM in a general case is best handled by radiative transfer simulations. However, such simulations can be even more challenging than those targeting the ionization field given that (i) the mean-free paths of relevant photons are longer, (ii) the frequency-dependence of the ionization and heating rates is important, which means multi-frequency calculations are necessary, and (iii) heating generally precedes reionization, meaning smaller halos must be resolved at earlier times.

It is useful to consider first a case in which heating of the IGM is spatially uniform, which could occur if the sources of heating have very hard spectra. In this limit, we can consider the evolution of the average background intensity,

Show simple models a la Pritchard & Loeb.

Spatial Structure

Talk about Jonathan's 2007 approach, Janakee's stuff, 21CMFAST approach, progress in RT sims (hard because X-ray mfp long). Ross et al. simulations.

Show example temperature power spectrum from, e.g., Pritchard & Furlanetto, and/or pictures from Andrei's 2013 paper. Point out dependence on normalization and spectral slope.

2.2.4   The Ly-α Background

This section needs quite a bit more work! Will discuss early Barkana stuff, Holzbauer & Furlanetto approach, semi-numeric treatments, Ahn et al.

Global Evolution

The Ly-α background intensity, which determines the strength of Wouthuysen-Field coupling [**?**, **?**], requires a special solution to the cosmological radiative transfer equation (see Eq. 2.20). Two effects separate this problem from the generic transfer problem outlined in the previous section: (i) the Lyman series forms a series of horizons for photons in the $10.2 < h\nu/\mathrm{eV} < 13.6$ interval, and (ii) the Ly-α background is sourced both by photons redshifting into the line resonance as well as those produced in cascades downward from higher n transitions.

It is customary to solve the RTE in small chunks in frequency space. Within each chunk, the optical depth of the IGM is small[8], while the edges are semi-permeable. For illustrative purposes, let us isolate the Ly-α background intensity sourced by photons redshifting into resonance from frequencies redward of Ly-β.

is computed analogously via

$$\widehat{J}_\alpha(z) = \frac{c}{4\pi}(1+z)^2 \sum_{n=2}^{n_{\max}} f^n_{\mathrm{rec}} \int_z^{z^{(n)}_{\max}} \frac{\varepsilon'_\nu(z')}{H(z')} dz' \qquad (2.42)$$

where f^n_{rec} is the "recycling fraction," that is, the fraction of photons that redshift into a Ly-n resonance that ultimately cascade through the Ly-α resonance [**?**]. We truncate the sum over Ly-n levels at $n_{\max} = 23$ as in [**?**], and neglect absorption by intergalactic H_2. The upper bound of the definite integral,

$$1 + z^{(n)}_{\max} = (1+z)\frac{\left[1 - (n+1)^{-2}\right]}{1 - n^{-2}}, \qquad (2.43)$$

is set by the horizon of Ly-n photons – a photon redshifting through the Ly-n resonance at z could only have been emitted at $z' < z^{(n)}_{\max}$, since emission at slightly higher redshift would mean the photon redshifted through the Ly-$(n+1)$ resonance.

Talk about excitation of Lyman alpha by photo-electrons.

---

[8]But for a small H_2 contribution, which here we neglect.

Spatial Fluctuations in the Ly-α background

Show a figure or two from Barkana & Loeb or Holzbauer & Furlanetto.

Ly-α Heating

Talk briefly about initial papers about Ly-α heating, the subsequent revisions, and the revival of this concept in the last year or so.

Excitation of Ly-α via fast photo-electrons

Brief discussion.

2.3 Sources of the UV and X-ray Background

In the previous section we outlined a procedure for evolving the ionization and temperature field without actually specificying the sources of ionization and heating. Instead, we used a generic emissivity, ε_ν, to encode the integrated emissions of sources at frequency ν within some region R, which we will now write as an integral over the differential luminosity function of sources, i.e.,

$$\varepsilon_\nu(z,R) = \int_0^\infty dL_\nu \frac{dn}{dL_\nu}. \tag{2.44}$$

Given the success of models which link the evolution of galaxies to the evolution of their host dark matter halos (citations), it is common to rewrite the emissivity as an integral over the DM halo mass function (HMF), dn/dm, multiplied by a conversion factor between halo mass and galaxy light, dm/dL_ν, i.e.,

$$\varepsilon_\nu(z,R) = \int_{m_{\min}}^\infty dm \frac{dn}{dm} \frac{dm}{dL_\nu}, \tag{2.45}$$

where $m_{\min}$ is the minimum mass of DM halos capable of hosting galaxies. Because dn/dm is reasonably well-determined from large N-body simulations of structure formation (citations), much of the modeling focus is on the mass-to-light ratio, dm/dL_ν, which encodes the efficiency with which galaxies form in halos and the relative luminosities of different kinds of sources within galaxies (e.g., stars, compact objects, diffuse gas) that emit at different frequencies[9].

The main strength of the 21-cm background as a probe of high-z galaxies is now apparent: though 21-cm measurements cannot constrain the properties of individual galaxies, they can constrain the properties of *all* galaxies, in aggregate, *even those too faint to be detected directly*. As a result, it is common to forego detailed modeling of the mass-to-light ratio and instead relate the emissivity to the fraction of mass in the Universe in collapsed objects,

$$\varepsilon_\nu(z,R) = \rho_b f_{\rm coll}(z,R) \zeta_\nu, \tag{2.46}$$

---

[9]Most models consider regions R that are sufficiently large that one can assume a well-populated HMF, though at very early times this approximation may break down, rendering stochasticity due to poor HMF sampling an important effect.

where the collapsed fraction is an integral over the HMF,

$$f_{\text{coll}} = \rho_m^{-1} \int_{m_{\min}}^{\infty} dm\, m \frac{dn}{dm} \tag{2.47}$$

and ζ_ν is an efficiency factor that quantifies the number of photons emitted at frequency ν per baryon of collapsed mass in the Universe. It is generally modeled as

$$\zeta_\nu = f_* N_\nu f_{\text{esc},\nu}, \tag{2.48}$$

where f_* is the star formation efficiency (SFE), N_ν is the number of photons emitted per stellar baryon at some frequency ν, and f_{esc} is the fraction of those photons that escape into the IGM. One could define additional ζ factors to represent, e.g., emission from black holes or exotic particles, in which case f_* and N_ν would be replaced by some black hole or exotic particle production efficiencies. In practice, three ζ factors are defined: ζ, ζ_X, and ζ_α, i.e., one efficiency factor for each radiation background of interest. A minimal model for the 21-cm background thus contains four parameters: $m_{\min}$, ζ, ζ_X, and ζ_α.

Because the factors within ζ are degenerate with each other, at least as far as 21-cm measurements are concerned, they generally are not treated separately as free parameters. However, it is still useful to consider each individually in order to determine a fiducial value of ζ and explore deviations from that fiducial model. In addition, inclusion of ancillary measurements may eventually allow ζ to be decomposed into its consituent parts. For the remainder of this section, we focus on plausible values of f_*, N_ν and f_{esc}.

2.3.1 Star Formation

Current high-z measurements support a relatively simple picture of star formation in early galaxies. The basic idea is that star formation is fueled by the inflow of gas from the IGM, but the overall rate of star formation in galaxies is self-limiting because winds and supernovae explosions expel gas that would otherwise form stars (many many references). Qualitatively, the need for some kind of feedback is apparent simply from the mismatch in shapes between the halo mass function and galaxy luminosity function, the latter of which is steeper at both the very bright and very faint ends.

Introduce MAR-driven models, semi-empirical constraints, etc.

Show example SFE recovered from UVLF fits. Variations in expectations at very high-z.

Pop III Star Formation

The very first generations of stars to form in the Universe did so under very different conditions than stars today, and it is not clear that the arguments outlined in the previous section apply. The first stars, by definition, formed from chemically-pristine material, since no previous generations of stars had existed to enrich the medium with heavy elements. This has long been recognized as a reason that the first stars are likely different than stars today (references). Without the energetically low-lying electronic transitions common in heavy elements, hydrogen-only gas clouds cannot cool efficiently, as collisions energetic enough to excite atoms from $n = 1$ to $n = 2$ (which subsequently cool via spontaneous emission of

Ly-α photons) imply temperatures of $\sim 10^4$ K. Halos with such virial temperatures are very rare at redshifts greater than $z \sim 10$.

However, other cooling channels may be available even in halos too small to support atomic (hydrogen) line cooling. Hydrogen molecules, H_2, can form using free electrons as a catalyst[10],

$$H + e^- \rightarrow H^- + h\nu \tag{2.49}$$

$$H^- + H \rightarrow H_2 + e^-, \tag{2.50}$$

These reactions are limited by the availability of free electrons[11] and the survivability of H^- ions. Even in the absence of astrophysical backgrounds, the formation of H_2 is limited by the CMB, which at the high redshifts of interest can dissociate the H^- ion. [?] found that the molecular hydrogen fraction in high-z halos scales with the virial temperature as

$$f_{H_2} \approx 3.5 \times 10^{-4} \left(\frac{T_{\mathrm{vir}}}{10^3 \ \mathrm{K}} \right)^{1.52}. \tag{2.51}$$

Once the first stars form, the situation grows considerably more complicated. As will be detailed in the following section (§2.3.2), massive stars are prodigious sources of UV photons. Some of these photons originate in the Lyman-Werner band (~ 11.2-13.6 eV), and are thus capable of dissociating molecular hydrogen. This processs is expected to quickly surpass H^- dissociation by the CMB as the most important mechanism capable of regulating star formation in chemically pristine halos.

A substantial literature has emerged in the last ~ 20 years aimed at understanding the critical LW background intensity, J_{LW}, required to prevent star formation in high-z mini-haloes.

$$M_{\mathrm{min}} = 2.5 \times 10^5 \left(\frac{1+z}{26} \right)^{-3/2} (1 + 6.96(4\pi J_{\mathrm{LW}})^{0.47}) \, M_\odot \tag{2.52}$$

Brief discussion of self-shielding complication.

Talk about metal enrichment and how this is (probably) the cause of the demise of PopIII.

Draw attention to how 21-cm background can contribute here: Ly-α and LW backgrounds are closely related. Note that there are very few studies of PopIII in the 21-cm background. Can point to Anastasia's stuff, my stuff, maybe Rick's stuff will be done soon...

Bottom line: m_{min} sort of understood, f_* unknown.

Show some models for the PopIII SFRD and compare to PopII SFRD extrapolated from observations and/or predictions from models.

2.3.2 UV Emission from Stars

Stellar photons are likely the dominant drivers of reionization[12] and the initial "activation" of the 21-cm background via Wouthuysen-Field coupling at $z \sim 30$. The 21-cm background

---

[10]Dust is the primary catalyst of H_2 formation in the local Universe, but of course is does not exist in the first collapsing clouds.

[11] Exotic models in which an X-ray background emerges before the formation of the first stars may similarly affect early star formation by boosting the electron fraction.

[12]There is still some room for a contribution from quasars [?, see, e.g.,]]Madau2018.

is thus sensitive to the spectral characteristics of stars in the Lyman continuum and Lyman Werner bands[13]. It is also in principle sensitive to the spectrum of even harder He-ionizing photons, since photo-electrons generated from helium ionization can heat and ionize the gas, while HeII recombinations can result in H-ionizing photons. The 21-cm signal could in principle even constrain the rest-frame infrared spectrum of stars in the early Universe, since IR photons can feedback on star-formation at very early times through H^- photo-detachment [?]. In this section, we focus only on the soft UV spectrum ($E < 54.4$ eV) to which the 21-cm background is most sensitive.

The most detailed predictions for stellar spectra come from stellar population synthesis (SPS) models, which take the following approach:

- Assume a model for the stellar initial mass function (IMF), $\xi(m)$, i.e., the number of relative number of stars formed in different mass bins. Commonly-adopted IMFs include Salpeter [?], Chabrier [?], Kroupa [?], and Scalo [?] which are all generally power-laws with indices ~ -2.3 , but differing in shape at the low mass end of the distribution ($M_* < 0.5\ M_\odot$).

- Assume a model for stellar evolution, i.e., how stars of different masses traverse the Hertzprung-Russell (HR) diagram over time.

- Assume a model for stellar atmospheres, i.e., as a function of stellar mass, age, and composition, determine the output spectrum.

With all these ingredients, one can synthesize a spectrum from a population of stars with a given age,

$$L_\nu(t) = \int_0^t dt' \int_{m_{\min}}^\infty dm \xi(m) l_\nu(m,t') \tag{2.53}$$

where $l_\nu(m,t)$ is the specific luminosity of a star of mass m and age t, and we have assumed that ξ is normalized to the mass of the star cluster, $\int dm \xi(m) = M_*$. Equation 2.53 can be generalized to determine the spectrum of a galaxy with an arbitrary star formation history (SFH) composed of discrete bursts. mention poor IMF sampling? Widely used stellar synthesis codes include STARBURST99 [?], BPASS [?], FSPS, Bruzual & Charlot...

Generally, 21-cm models do not operate at level of SPS models because the 21-cm background is insensitive to the detailed spectra and SFHs of individual galaxies. Instead, because 21-cm measurements probe the relatively narrow intervals $10.2 < h\nu/\mathrm{eV} < 13.6$ via Wouthuysen-Field coupling and $h\nu > 13.6$ eV through the ionization field, it is common to distill the predictions of SPS models into just two numbers, N_{ion} and N_α, which integrate over age and the details of the stellar SED, i.e.,

$$N_{\mathrm{ion}} = m_*^{-1} \int_0^\infty dt' \int_{\nu_{\mathrm{LL}}}^\infty \frac{d\nu}{h\nu} L_\nu(t') \tag{2.54}$$

$$N_\alpha = m_*^{-1} \int_0^\infty dt' \int_{\nu_\alpha}^{\nu_{\mathrm{LL}}} \frac{d\nu}{h\nu} L_\nu(t') \tag{2.55}$$

---

[13]We use this definition here loosely. Technically, the LW band is $\sim 11.2 - 13.6$ eV, a range which bounds photons capable of photo-dissociating molecular hydrogen, H_2. The Ly-α background is sourced by photons in a slightly broader interval, $\sim 10.2 - 13.6$ eV, but it is tedious to continually indicate this distinction, and as a result, we use "LW band" to mean all photons capable of eventually generating Ly-α photons.

where ν_{LL} is the frequency of the Lyman limit (13.6 eV) and ν_α is the Ly-α frequency. UV emission is dominated by massive, short-lived stars, hence the integration from $t = 0$ to $t = \infty$.

Assuming a Scalo IMF, stellar metallicity of $Z = Z_\odot/20$, using STARBURST99 SPS model, [?] report $N_\alpha = 9690$, further broken down into sub-intervals between each Ly-n resonance, an oft-used reference value even today. The canonical value of $N_{ion} = 4000$ (I think) makes the same assumptions but I can't find where this first appeared. The general expectation is for N_{ion} and N_α increase for more top-heavy IMF and lower metallicity, meaning these values are likely to increase for Pop III stars (citations). Similarly, binary evolution can effectively increase the lifetimes of massive stars, leading to a net gain in UV photon production [?].

Note that detailed SPS may be needed to if jointly fitting 21-cm measurements and galaxy population.

Show PopII and PopIII spectra?

2.3.3   Escape of UV Photons from Galaxies

Summarize briefly the status of f_{esc}.

2.3.4   X-rays from Black Holes

Though stars themselves emit few photons at energies above the HeII-ionizing edge (~ 54.4 eV), their remnants can be strong X-ray sources. While solitary remnants will be unlikely to accrete much gas from the diffuse ISM, remants in binary systems may accrete gas from their companions, either via Roche-lobe overflow or stellar winds. Such systems are known as X-ray binaries (XRBs), further categorized by the mass of the donor star: "low-mass X-ray binaries" (LMXBs) are those fueled by Roche-lobe overflow from a low-mass companion, while "high-mass X-ray binaries" (HMXBs) are fed by the winds of massive companions. XRBs exhibit a rich phenomenology of time- and frequency-dependent behavior and are thus interesting in their own right. For a review see, e.g., [?].

In nearby star-forming galaxies, the X-ray luminosity is generally dominated by the HMXBs [?, ?, ?]. Furthermore, the total luminosity in HMXBs scales with the star formation rate, as expected given that the donor stars in these systems are massive, short-lived stars. An oft-used result in the 21-cm literature stems from the work of [?] (update of Gilfanov), who find

$$L_X = 2.6 \times 10^{39} \left(\frac{\dot{M}_*}{M_\odot \text{ yr}^{-1}} \right) \text{ erg s}^{-1} \qquad (2.56)$$

where L_X refers to the 0.5-8 keV band. This relation provides an initial guess for many 21-cm models, which add an extra factor f_X to parameterize our ignorance of how this relation evolves with cosmic time. For example, [?] write

$$L_X = 3 \times 10^{40} f_X \left(\frac{\dot{M}_*}{M_\odot \text{ yr}^{-1}} \right) \text{ erg s}^{-1}, \qquad (2.57)$$

which is simply Equation 2.56 re-normalized to a broader energy range, $0.2 < h\nu/\text{keV} < 3 \times 10^4$, assuming a power-law spectrum with spectral index $\alpha_X = -1.5$, where α_X is defined by $L_E \propto E^{\alpha_X}$, with L_E in energy units.

The normalization of these empirical L_X-SFR relations are not entirely unexpected, at least at the order-of-magnitude level. For example, if one considers a galaxy forming stars at a constant rate, a fraction $f_\bullet \simeq 10^{-3}$ of stars will be massive enough ($M_* > 20 \, M_\odot$) to form a black hole assuming a Chabrier IMF. Of those, a fraction f_{bin} will have binary companions, with a fraction f_{surv} surviving the explosion of the first star for a time τ. If accretion onto these black holes occurs in an optically thin, geometrically-thin disk with radiative efficiency $\varepsilon_\bullet = 0.1$ which obeys the Eddington limit, then a multi-color disk spectrum is appropriate and a fraction $f_{0.5-8} = 0.84$ of the bolometric luminosity will originate in the 0.5-8 keV band. Finally, assuming these BHs are "active" for a fraction f_{act} of the time, we can write [?, ?]

$$L_X \sim 2 \times 10^{39} \text{erg s}^{-1} \left(\frac{\dot{M}_*}{M_\odot \text{ s}^{-1}} \right) \left(\frac{\varepsilon_\bullet}{0.1} \right) \left(\frac{f_\bullet}{10^{-3}} \right) \left(\frac{f_{\text{bin}}}{0.5} \right) \left(\frac{f_{\text{surv}}}{0.2} \right) \left(\frac{\tau}{20 \text{ Myr}} \right) \left(\frac{f_{\text{act}}}{0.1} \right) \left(\frac{f_{0.5-8}}{0.84} \right).$$
$$(2.58)$$

While several of these factors are uncertain, particularly f_{surv} and f_{act}, this expression provides useful guidance in setting expectations for high redshift. For example, it has long been predicted that the first generations of stars were more massive on average than stars today owing to inefficient cooling in their birth clouds. This would boost $f_\bullet$, and thus $L_X/\dot{M}_*$, so long as most stars are not in the pair-instability supernova (PISN) mass range, in which no remnants are expected.

There are of course additional arguments not present in Eq. 2.58. For example, the MCD spectrum is only a good representation of HMXB spectra in the "high soft" state. At other times, in the so-called "low hard" state, HMXB spectra are well fit by a power-law. The relative amount of time spent in each of these states is unknown.

In addition, physical models for the L_X-SFR relation may invoke the metallicity as a driver of changes in the relation with time and/or galaxy (stellar) mass. As the metallicity declines, one might expect the stellar IMF to change (as outlined above), however, the winds of massive stars responsible for transferring material to BHs will also grow weaker as the opacity of their atmospheres decline. As a result, increases in L_X/SFR likely saturate below some critical metallicity. Observations of nearby, metal-poor dwarf galaxies support this picture, with L_X/SFR reaching a maximum of ~ 10 times the canonical relation quoted in Eq. 2.56 [?].

Show different spectra often adopted in models and discuss.

Super-Massive Black Holes

Say a few words about plausibility, reference [?].

2.3.5   X-rays from Shocks and Hot Gas

While compact remnants of massive stars are likely the leading producer of X-rays in high-z star-forming galaxies (see previous sub-section), the supernovae events in which these objects are formed may not be far behind. Supernovae inject a tremendous amount of energy

into the surrounding medium, which then cools either via inverse Compton emission (in supernova remnants; [**?**]) or eventually via bremsstrahling radiation (in the hot interstellar medium; ISM. Because these sources are related to the deaths of massive stars their luminosity is expected to scale with SFR, as in the case of HMXBs. Indeed, [**?**] find that diffuse X-ray emission in nearby sources follows the following relation in the 0.5-2 keV band:

$$L_X = 8.3 \times 10^{38} \left(\frac{\dot{M}_*}{M_\odot \ \mathrm{yr}^{-1}} \right) \ \mathrm{erg \ s}^{-1} \tag{2.59}$$

This luminosity is that from all unresolved emission, and as a result, is not expected to trace emission from the hot ISM alone. Emission from supernova remnants will also contribute to this luminosity, as will fainter, unresolved HMXBs and LMXBs. [**?**] estimate that $\sim 30 - 40\%$ of this emission may be due to unresolved point sources.

Though the soft X-ray luminosity from hot gas appears to be subdominant to the HMXB component in nearby galaxies, there are of course uncertainties in how these relations evolve. Furthermore, the bremmstrahlung emission characteristic of hot ISM gas has a much steeper $\sim \nu^{-2.5}$ spectrum than inverse Compton ($\sim \nu^{-1}$) or XRBs ($\sim \nu^{-1}$ or $\nu^{-1.5}$), and thus may heat more efficiently (owing to $\sigma \propto \nu^{-3}$ cross section) provided soft X-rays can escape galaxies.

2.3.6 Escape of X-rays from Galaxies

Though the mean free paths of X-rays are longer than those of UV photons, they still may not all escape from galaxies into the IGM.

The column density is used to account for absorption by neutral hydrogen in the ISM, which hardens the intrinsic spectrum. Simulations suggest typical values of $N_{H\mathrm{I}} \sim 10^{21}$ cm^{-2} [**?**], which is substantial enough to eliminate emission below ~ 0.5 keV.

Given the many unknowns regarding X-ray emission in the early Universe, 21-cm models often employ a three-parameter approach, i.e., instead of a single value of ζ_X, the specific X-ray luminosity is modeled as

$$L_{X,\nu} = L_{X,0} \left(\frac{h\nu}{1\mathrm{keV}} \right)^{\alpha_X} \exp\left[-\sigma_\nu N_{H\mathrm{I}}\right] \tag{2.60}$$

and the normalization, $L_{X,0}$, spectral index α_X, and typical column density, $N_{H\mathrm{I}}$, are left as free parameters.

It is common to approximate this intrinsic attenuation with a piecewise model for L_X, i.e.,

$$L_{X,\nu} = \begin{cases} 0 & h\nu < E_{\min} \\ L_{X,0} \left(\frac{h\nu}{1\mathrm{keV}} \right)^{\alpha_X} & h\nu \geq E_{\min} \end{cases} \tag{2.61}$$

Note that $N_{H\mathrm{I}}$ (or $E_{\min}$) can be degenerate with the intrinsic spectrum, e.g., the SED of HMXBs in the high-soft state exhibits a turn-over at energies $h\nu < 1$ keV, which could be mistaken for strong intrinsic absorption.

Translate this to ζ_X.

2.3.7 Cosmic Rays from Supernovae

Other sources of high energy radiation have been explored in recent years though are generally found to be sub-dominant. However, surprises may be in store...

2.4 Predictions for the 21-cm Background

Over the last four sections we have assembled a simple physical picture of the IGM at high redshift from which we can derive predictions for the 21-cm brightness temperature. Here, we finally describe the generic sequence of events predicted in most models, and the sensitivity of the 21-cm background to various model parameters of interest.

2.4.1 Generic Series of Events

Figure XYZ depicts what are now standard predictions for the global 21-cm signal (top) and power spectrum (bottom). Time proceeds from left to right from the Big Bang until the end of reionization. There are four distinct epochs within this time period, labeled A, B, C, and D, which we describe in more detail below.

A. The Dark Ages: As the Universe expands after cosmological recombination, Compton scattering between free electrons and photons keep the radiation and matter temperature in equilibrium. The density is high enough the collisional coupling remains effective, and so $T_S = T_K = T_{CMB}$. Eventually, Compton scattering becomes inefficient as the CMB cools and the density continues to fall, which allows the gas to cool faster than the CMB (see also earlier figures). Collisional coupling remains effective for a short time longer and so T_K follows T_S. This results in the first decoupling of T_S from T_{CMB} at $z \sim 80$, and thus an absorption signature at $\nu \sim 15$ MHz, which comes to an end as collisional coupling becomes inefficient, leaving T_S to reflect T_{CMB} once again.

B. First Light: When the first stars form they flood the IGM with UV photons for the first time. While Lyman continuum photons are trapped near sources, photons with energies $10.2 < h\nu/\text{eV} < 13.6$ either redshift directly through the Ly-α resonance or cascade via higher Ly-n levels, giving rise to a large-scale Ly-α background capable of triggeiring Wouthuysen-Field coupling as they scatter through the medium. As a result, T_S is driven back toward T_K, which (in most models) still reflects the cold temperatures of an adiabatically-cooling IGM.

C. X-ray Heating: The first generations of stars beget the first generations of X-ray sources, whether they be the explosions of the first stars themselves or remnant neutron stars or black holes that subsequently accrete. Though the details change depending on the identity of the first X-ray sources, generally such sources provide photons energetic enough to travel great distances. Upon absorption, they heat and partially ionize the gas, eventually driving $T_S > T_{CMB}$. Once $T_S \gg T_{CMB}$, the 21-cm signal "saturates," and subsequently sensitive only to the density and ionization fields. However, it is possible that heating is never "complete" in this sense before the completion of reionization,

name	description	typical values
ζ_i	Ionizing photon production efficiency	40 ish
ζ_α	Ly-α photon production efficiency	40 ish
ζ_X	X-ray photon production efficiency	xxx
$T_{\min}$	Minimum virial temperature of star-forming halos	10^4 K

Table 2.1: Parameters in simple 21-cm models.

meaning neutral pockets of IGM gas may remain at temperatures at or below $T_{\rm CMB}$ until they are engulfed in the overlap event of large ionized bubbles.

D. Reionization: As the global star formation rate density climbs, the growth of ionized regions around groups and clusters of galaxies will continue, eventually culminating in the completion of cosmic reionization. This rise in ionization corresponds to a decline in the amount of neutral hydrogen in the Universe capable of producing or absorbing 21-cm radiation. As a result, the amplitude of the 21-cm signal, both in its mean and fluctuations, falls as reionization progresses.

Draw connection between features in 21-cm signal and evolution of ionization, heating, and J_α introduced in earlier sections.

The evolution of 21-cm fluctuations is more complicated, though this same series of events imprints on fluctuation patterns as well.

2.4.2 Sensitivity to Model Parameters

Use this section to highlight the sensitivity to parameters in more detail.

Things to discuss:

- Sensitivity to $T_{\min}$ and f_*.

- Sensitivity to α_X, $E_{\min}$, and f_X.

- Sensitivity to ζ and $R_{\rm mfp}$ (and updates)

- Clumping, feedback

- Shot noise in galaxy counts in voxels.

- PopIII stuff. AGN stuff.

- Exotic physics? Defer to Jonathan's chapter.

2.4.3 Modeling Tools

Predictions from previous section came from a mix of different groups and codes. Discuss some differences here.

- 21CMFAST and DEXM

- Anastasia's code

- simfast21

- ARES

- RT simulations

Chapter 3

Physical cosmology from the 21-cm line

Jonathan Pritchard

a discussion of using 21-cm intensity mapping/interferometry to constrain the matter power spectrum and cosmological parameters. The signal can also be used as a signpost for new physics beyond the standard model, such as heating and cooling of the primordial gas through interactions with dark matter. Subtopics could include: redshift space distortions and their likely observability, measuring the matter power spectrum during the cosmic dawn and (eventually) the dark ages, IGM heating through dark matter annihilations, IGM cooling through interactions with exotic dark matter candidates.

This chapter discusses some important things

3.1 Physical cosmology through the 21-cm epoch

Summarise the different ways in which cosmology underpins the nature of the 21 cm signal.

3.1.1 Dark matter

Discuss different dark matter models and impact

3.1.2 PMBH

Discuss primordial black holes and other exotic objects

3.2 Unique signatures of cosmology

What might be distinctive signatures of cosmology in the 21cm signal?

3.2.1   Non-Gaussianity

Impact of clustering on bubbles.

3.2.2   IGM Heating

Various ways exotic physics can impact 21 cm signal

3.3   Separating cosmology from astrophysics

How might cosmology be separated from astrophysics?

3.3.1   Redshift space distortions

As route to separating different elements of signal

3.3.2   Time evolution

Looking at the dark ages, or via time evolution, or going to late times to post-EoR.

3.4   Summary

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis eu egestas erat. Maecenas tincidunt lacinia tincidunt. Mauris id lectus nec neque feugiat condimentum vitae at diam. In vel orci nunc, non commodo mauris. Vivamus ipsum enim, vulputate quis pharetra non, molestie quis felis. Vivamus porttitor placerat turpis at accumsan. Nunc tortor velit, faucibus a rhoncus nec, blandit non elit. Nam consectetur lectus eu nisi blandit dapibus rhoncus dui tempus. Mauris fermentum dolor vel ipsum vulputate sit amet ultricies tortor lacinia. Donec ut nibh erat. Morbi nec mi ante. Integer nec vestibulum diam. Donec tincidunt pellentesque quam, ut interdum mauris venenatis condimentum. Nam condimentum, augue in aliquet gravida, neque dui elementum eros, id semper eros purus sed felis. Curabitur in justo sit amet sapien ultrices hendrerit at quis nibh. Quisque iaculis pulvinar tincidunt.

$$
\begin{aligned}
C(12) &= \left[\vec{\pi} \cdot \vec{\phi}\,(x+r) \right] \\
&\approx 1 - \text{const} \frac{r^2}{L^2} \int_r^L \frac{x\mathrm{d}x}{x^2} + \cdots \\
&\approx 1 - \text{const} \frac{r^2}{L^2} \ln \frac{x\mathrm{d}x}{x^2} + \cdots .
\end{aligned}
\tag{3.1}
$$

Aenean tellus risus, porta sit amet porta vitae, tincidunt ut felis. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Phasellus pulvinar placerat velit auctor egestas. Vivamus euismod fringilla tincidunt. Sed ut magna felis, id sollicitudin nunc. Quisque a dui eu erat consectetur egestas a quis justo. Aenean euismod congue diam,

Figure 3.1: This is figure 1 in chapter 1.

Table 3.1: Greek Letters.

α	β	γ	δ	ε	ϵ	ζ	η
θ	ϑ	γ	κ	λ	μ	ν	ξ
o	π	ϖ	ρ	ϱ	σ	ς	
τ	υ	ϕ	φ	χ	ψ	ω	
Γ	Δ	Θ	Λ	Ξ	Π	Σ	Υ
Φ	Ψ	Ω					

vel posuere urna fermentum sit amet. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris faucibus lacus eget est mollis auctor. Donec at nibh ligula, et posuere massa. Phasellus quis leo diam [**?**]. Donec aliquam blandit risus, eu venenatis ante euismod eu. Curabitur cursus justo id arcu condimentum feugiat. Integer sapien urna, vulputate et adipiscing nec, convallis et justo. Suspendisse in ipsum at felis ornare interdum [**?**],

Cras adipiscing sagittis nunc vel luctus. Suspendisse volutpat augue quis erat semper consequat dignissim tellus euismod. Morbi hendrerit, tellus id aliquam iaculis, nibh leo tincidunt eros, vitae varius ligula felis in mi.
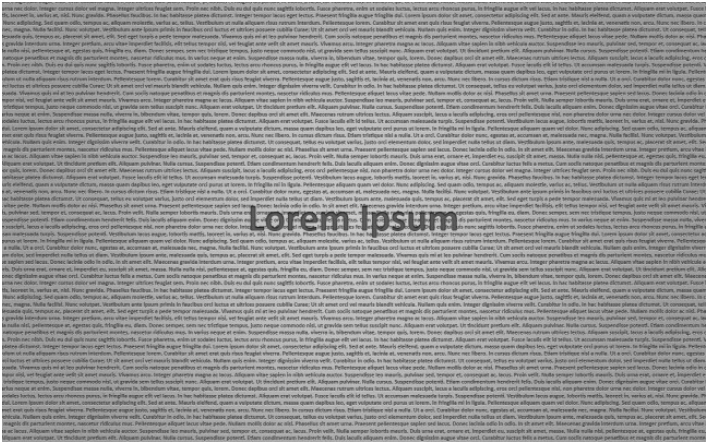
Figure 3.2: This is figure 2 in chapter 1.

Chapter 4

Inference from the 21cm signal

Bradley Greig

Abstract

The 21cm signal is physics rich and encodes a wealth of astrophysical and cosmological information (refer back to previous chapters here). However, how do we extract this information? This chapter focusses on how we intend to extract the interesting physics from the observed 21cm signal.

Essentially, the problem can be broken down into three parts:

1. Compressing the observed data into a manageable dataset (i.e. statistic or images) to tease out the interesting physics. Which is the optimal statistic, and what in particular is it optimal for?

2. An efficient means to model/simulate the 21cm signal

3. A robust probabilistic framework to compare simulations to observations to extract the physics.

In this chapter we will focus on each separately, before concluding with the current state of the art in inference of the 21cm signal.

4.1   What do we actually measure?

Ultimately, we observe the brightness temperature fluctuations in the 21cm signal, relative to the background CMB temperature. In effect, we observe $\delta T_{b,21\mathrm{cm}}(\nu, \mathbf{x})$, dependant on the observed frequency (line of sight distance) and the 2D angular position on the sky. Interferometric experiments (e.g.) enable us to access the full 3D information of the 21cm signal. Global experiments on the other hand average out the brightness temperature over the entire sky, compressing the information into a single average brightness as a function of frequency.

Once we have a measurement of this 3D 21cm signal, we are free to compress the data in an innumerable number of ways. We can extract single 2D tomographic images of the differential brightness temperature as a function of redshift (frequency) enabling us to be able

to directly observe the topology of the neutral and ionised gas during reionisation. Alternatively we can average out the 3D information either spatially or in frequency to maximise the signal to noise in a certain regime to maximise certain astrophysical processes and features expected during reionisation.

4.2 Optimal methods for extracting the astrophysics and cosmology from the 21cm signal

To infer the interesting astrophysics and cosmology from the observational data, we need a way to characterise the spatially and time varying 21cm signal. The signal is weak, therefore we likely will need to compress the data in some way to tease out the interesting physics from the noise.

Here, list a variety of known ways in which can be used to extract information about the 21cm signal. This will include both imaging and statistical techniques. For each, provide a brief discussion about the method (i.e. what it is) and then what leverage it provides in regard to the astrophysics etc. and what specific features are useful **(this may marginally overlap with Jordan's section on Predictions and Challenges)**. At various points in this section, can refer back to previous chapters when specific physics is being referred to. For example, if a specific statistic is sensitive to a certain population of sources or an epoch of history, refer back to the appropriate section in the previous chapters.

This should be a considerable more in-depth discussion that what is covered in Jordan's and Jonathan Pritchard's chapters. Need to gauge the overlap between these chapters.

Need to work in Adrian Liu's paper highlighting the importance of knowing τ for precision recovery of cosmology

4.2.1 Minkowski functionals

Compresses the information down into a single number, which can describe the connectedness and overall topology of the 21cm signal. Cite Zaroubi, also a recent arXiv paper (Chen maybe).

4.2.2 Global signal

Describes the entire globally averaged history of the 21cm signal. Turning points/features within the signal can be used to disentangle information about sources etc. Can also throw in some discussion regarding EDGES/DM, i.e. fundamental cosmology + astrophysics. **How much overlap is there going to be between this and Jordan/Jonathan Pritchard's chapters.?**

There are a variety of papers investigating the global signal. Could be a few different ways to use the signal to extract information. **This may overlap with Jordan's chapter on predictions and challenges.**

4.2.3 Power spectrum

The main goto statistic to describe the signal. Sensitive to a vast array of astrophysics. **This may overlap with Jordan's chapter on predictions and challenges.**

However, doesn't fully describe the signal. Power spectrum is a Gaussian statistic, the signal is inherently non-Gaussian. Therefore information loss.

This should focus only on the 3D spherically averaged 21cm PS. Can mention that the 2D cylindrically averaged PS can in principle also be used, and why it would be useful.

4.2.4 Skewness/Kurtosis/PDFs etc.

A means to measure the non-Gaussianity of the signal, however, compresses the information as it is a scale independent quantity. Useful for describing topology and evolution of the signal. Has certain characteristic features of interest. See Catherine's paper and others.

4.2.5 Bispectrum

A scale dependant measure of the non-Gaussianity of the 21cm signal. Superior to both the power spectrum and skewness etc.. More complicated to visualise/understand. Has been used to describe the EoR quite extensively. Discuss the variety of configurations and their utility/usefulness. See Suman's, and other publications.

4.2.6 Wavelets

An alternative to simply Fourier transforming the data, a family of wavelets can be used to transform the 21cm data to a more useful basis. For example, the Morlet transform allows the transform to be localised both in frequency and in space, improving over the Fourier transform which can only be localised in one or the other.

In theory, it should contain similar, if not more information that the power spectrum, as it can disentangle line-of-sight effects better. See Cath's paper.

4.2.7 Individual bubbles

Images will provide a direct tangible link to the process of reionisation. Revealing exactly where ionising bubbles are, and thus where to look for the sources responsible for the bubble creation.

However, bubble identification will become rather problematic, as it is the differential brightness that is observed, not the raw brightness temperature. Need smart/sophisticated approaches to search and characterise the signal.

Look at individual regions of interest, i.e. around a bright QSO or large number of galaxies.

Matched filters etc., are useful for finding/detecting isolated bubbles

4.2.8 Stacked images

May be difficult to detect individual objects, instead, one could stack 21cm spectra centred on known galaxies. See Paul Geil's paper.

Border's on potential discussions of synergies. Requires known locations of ionising sources with precise redshifts and positions of the sky. JWST/WFIRST. Is there

4.2.9 Bubble size distributions/statistics

Constructing a statistical distribution of the ionised bubbles through redshift can reveal information about the size and number of galaxies responsible for reionisation (also the production rate of the ionising photons).

Techniques to recover statistical distributions are for example Sambit's super pixels stuff. Need to look at Koki's granulometry stuff to see where that sits in this context.

4.2.10 Other statistics

Are there other statistics that I have overlooked/forgotten. Need to do a search of the literature to ensure I have covered everything.

4.3 Efficient methods to model/simulate the 21cm signal

Ideally we want to use the largest, most physically accurate simulations to match the observed 21cm signal. However, this is not practical. Instead, we must come up with methods to compensate accuracy for efficiency.

Originally I envisaged this section to go as follows

- Discuss briefly that numerical simulations are great but too computationally expensive

- Highlight semi-numerical/analytic simulations as fitting the bill by being faster etc.

- Then discuss alternatives (i.e. emulators).

This will need to be re-worked given Jordan's plan to discuss this. I haven't yet thought of a logical plan to motivate this. I guess one way would be to concatenate it into just a simulation section and briefly summarise what was discussed in Jordan's chapter (which is a couple of chapters ago, so might be appropriate to do so). It's less obvious to move into discussing emulators in the absence of motivating the need for them by highlighting the complications of simulations in general.

4.3.1 Numerical simulations

Ideally, use large, numerical simulations to make a realisation which matches observation. Want to include as much physics as possible into these simulations. Doing so, comes at a serious cost. Simulating the 21cm signal is complicated! Brief summary on the required

dynamic range. Briefly highlight the expensive nature of these simulations, Hydrodynamics, radiative transfer etc. Quickly becomes computationally prohibitive to run more than a couple of realisations. **Jordan's section (Modelling Tools) will likely cover most if not all of this**.

4.3.2 Semi-numerical/analytic models of the 21cm signal

Enter semi-numerical approaches, which broadly encapsulate the global average quantities of reionisation and the cosmic dawn and reproduce morphologically similar realisations of the 3D structure. The advantage here is that the computational costs are drastically reduced (orders of magnitude less). **Jordan's section (Modelling Tools) will likely cover most if not all of this**.

4.3.3 Intelligent sampling of the parameter space

Running simulations to span all of the allowed parameter space may be too computationally intensive. However, perhaps we can make intelligent assumptions/guesses about how many simulations to run, and on a specific area of physics to focus on. Some of the concepts here may overlap slightly with the Emulator section below. This sampling was discussed in one of Benoit's recent papers.

4.3.4 Emulators

An alternative to directly running a simulation to estimate some astrophysical statistic/model, one can instead construct a function which estimates what the statistical signal should be, given some astrophysical or cosmological parameter set. This is what is referred to as an emulator. Using simulations, a generator function is constructed which can approximate the statistics of the signal. Results in multiple orders of magnitude improvement in computational speed as no new simulations are required. However, it can only approximate statistics, it does not generate 3D simulations.

Developing an emulator benefits from efficient parameter space approaches, as it minimises the size of the database required to construct the emulator. There are a variety of approaches to consider when attempting to minimise the sampling of the astrophysical parameter space. For example Latine-Hypercube, Gaussian processes... Can briefly discuss each method, and a few recent papers that explore methods to do this

There are numerous machine learning approaches to construct an emulator.

Discuss Nick Kern's of 21cmFAST and Cohen's global signal one of Anastasia's code.

4.3.5 Characterising our ignorance

We are fully aware that semi-numerical approaches are inaccurate at 10s of per cent level. Additionally, they oversimplify/completely ignore the underlying astrophysical properties. However, if the global quantities (observables such as luminosity functions etc.) can still be reproduced within agreement, we can develop an understanding of how one might map

from a semi-numerical simulation, to a more realistic simulation. Basically, tell the large, computationally expensive simulation where exactly to look in the region of parameter space.

Using summary statistics and globally average observables, we can develop a means to calibrate one simulation to mimic the outputs of another. In other words, develop a bias or functional form to smooth out uncertainties from one simulation to mimic the results of a more detailed simulation.

Describe ongoing attempts to quantify this. i.e. use luminosity functions to calibrate simulations, apply redshift corrections to deal with photon non-conservation etc.

These correction factors etc., will be crucial for inferring astrophysics and cosmology from the 21cm signal

4.4   Inference methods for 21cm

Having discussed methods to model/simulate the 21cm signal, now need to shift focus to methods to infer information about the astrophysics/cosmology from the 21cm signal.

4.4.1   Fisher Matrices

Simplest method, which takes derivatives with respect to model parameters of a functional form (i.e. a 21cm statistical signal) to infer parameter constraints. Effectively, exploits how sensitive the 21cm signal is to specific parameters. Limiting assumption include Gaussianity etc.

Successfully used for cosmology

4.4.2   Bayesian MCMC

Significantly more robust method to parameter inference. Outline Bayes' theorem. Highlight that this basically boils down to exploring by random walks, using a likelihood function and priors to accept/penalise regions of parameter space. Very useful for recovering constraints. Again, successfully used in Cosmology etc.

Describe its use in the context of 21cm. Introduce 21CMMC, and a couple of other codes that use MCMC (Sultan's simfast21, Jordan's global signal, Geraint's MCMC, Simon's Mhysa). Highlight that 21CMMC is the only direct MCMC. Emphasise that techniques such as emulators can be coupled with MCMC to improve overall computational efficiency at the cost of some further inaccuracies.

4.4.3   Nested sampling and model inference

We have a wide variety of simulations, each with their own strength's and weaknesses. In principle, various models/simulations/physical prescriptions can be discriminated against using model selection or inference. Related to MCMC, nested sampling can focus on model selection rather than simply astrophysical recovery. Can refer to Tom Binnie's recent paper.

4.4.4 Neural Networks

Instead of using MCMC to recover parameters from statistics, we can use the full 2/3D images of the 21cm signal with neural networks. One can construct a database of simulations and extract the 21cm signal to construct a neural network. This network can then learn an large number of properties, i.e. how changes in the topology are affected by astrophysical parameters etc. Using this technique one can directly convert from an observed 2D image to the underlying astrophysics. Like emulators, the application of the network is extremely quick. Downsides are parameter errors etc. and in some sense physical intuition. It tells you an answer, but not why.

Simulated images of the 21cm signal have already been used to infer astrophysical and cosmological constraints. For example with convolutional networks (e.g Nicolas or Paul La Plante). Discuss other machine learning approaches.

Bibliography

Chapter 5

Observational strategies: power spectra and images

Gianni Bernardi (INAF-IRA)

Abstract

This chapter reviews the basics of radio interferometry (van-Citter-Zernike theorem, uv-coverage, image formation calibration) and how they are linked to the measurements of the 21-cm power spectrum and its tomography

5.1   Chapter layout

- review of basics of interferometry: van-Citter-Zernike theorem, *uv*-coverage, image formation;

- from visibilities to power spectra (in particular the delay transform approach - some overlap with Trott's chapter);

- interferometric calibration: impact on foreground subtraction, calibration errors redundant calibration;

- array design: power spectrum vs imaging, minimmum vs maximum redundancy - some overlap with Trott's chapter?;

- impact of calibration errors on power spectra;

- ionospheric impact/calibration;

- 21-cm image tomography;

5.2   Interferometry overview

The Van Cittert-Zernike theorem expresses the fundamental relationship between the sky spatial brightness (or brightness distribution) I and the quantity measured by an interferometer, i.e. the visibility V (e.g., [1]):

$$V_{ij}(\mathbf{b}, \lambda) = \int_\Omega \bar{I}(\hat{\sigma}, \lambda) e^{-2\pi i \mathbf{b} \cdot \hat{\sigma}} d\sigma, \tag{5.1}$$

where $\mathbf{b}$ is the baseline vector that separates between antenna i and j and $\hat{sigma}$ is the observing direction (see Figure ???). The baseline vector is here specified in wevelengths, i.e. $\mathbf{b} = \frac{\mathbf{b}_m}{\lambda}$, where $\mathbf{b}_m$ is the baseline vector expressed in meters and λ is the observing wavelength. The sky brightness distribution does not enter directly in equation 5.1, but filtered by the antenna primary beam response A that depends upon the direction in the sky and the wavelength, i.e. $\bar{I} = AI$. The response of the primary beam attenuates the sky emission away from the pointing direction, effectively reducing the field of view θ of the instrument. Generally speaking, the size of the field of view is essentially given by the antenna diamater D:

$$\theta \approx \frac{\lambda}{D}. \tag{5.2}$$

The integral of equation 5.1 is taken over the source size Ω. Equation 5.1 is often re-written in a different coordinate system, i.e. the components of the baseline vector (u, v, w) and the reciprocal (l, m, n), where (l, m) are the coordinates in the plane on the sky tangent to the observing direction n (for a detailed discussion on coordinate systems see, for example [1]). Using this different coordinate system, equation 5.1 becomes (e.g., [1]):

$$V_{ij}(u, v, w, \lambda) = \int_\Omega \bar{I}(\hat{\sigma}, \lambda) e^{-2\pi i(ul + vm + wn)} dl\, dm\, dn, \tag{5.3}$$

(GB: fix the above equation)...

   Although low frequency radio observations are intrinsically wide-field, for the purpose of studying the 21 cm observables, we can reduce equation 5.3 to a two dimensional Fourier transform:

$$V_{ij}(u, v, \lambda) = \int_\Omega \bar{I}(l, m, \lambda) e^{-2\pi i(ul + vm)} dl\, dm. \tag{5.4}$$

Equation 5.4 indicates that an *interferometer measures the two dimensional Fourier transform of the spatial sky brightness distribution*. If our goal is to reconstruct the sky brightness distribution, equation 5.4 can be inverted into its corresponding Fourier pair:

$$\bar{I}(l, m, \lambda) = \int_{-\infty}^{+\infty} V_{ij}(u, v, \lambda) e^{2\pi i(ul + vm)} du\, dv. \tag{5.5}$$

Equation 5.5 is, however, a poor reconstruction of the sky brightness distribution as only one Fourier mode is sampled at the time t. Strictly speaking, indeed, all the quantities in equation 5.4 and 5.5 are time variable. In most cases, the time dependence of the primary beam and the sky brightness distribution can be neglected, however, this is not the case for the visibility V as the projection of the baseline vector with respect to the source direction

changes significantly throughout a long (e.g. a few hours) track. In this way, many measurements of the visibility coherence function V as $(u, v,)$ change with time can be made, allowing for a better reconstruction of the $\bar{I}(l, m, \lambda)$ function. This methods is commonly described as *filling the uv plane via Earth rotation synthesis* and was invented by [**?**]. The other (complementary) way to fill the *uv* plane is to deploy more antennas on the ground in order to increase the number of instantaneous measurements of independent Fourier modes. If N antennas are connected in an interferometric array, $\frac{N(N-1)}{2}$ instantaneous measurements are made.

The combination of a large number of antennas and the Earth rotation synthesis, defines the sampling function $S(u, v,)$ in the *uv* plane. In any real case, equation 5.5 can therefore be re-written as:

$$\bar{I}_D(u, v, \lambda) = \int_{-\infty}^{+\infty} S(u, v,) V(u, v, \lambda) e^{2\pi i (ul + vm)} du \, dv \qquad (5.6)$$

where $\tilde{I}_D$ indicates the sky brightness distribution sampled at a finite number of (u, v) points (often termed *dirty image*) and I dropped the explicit dependence on the antenna pair as redundant at his point. Using the convolution theorem, equation 5.7 can be re-written as:

$$\bar{I}_D = \tilde{S}V = \tilde{S} * \tilde{V} = \text{PSF} * \tilde{V}. \qquad (5.7)$$

where

I will give some examples of sampling functions for different instruments in Section 5.4, however, the sampling function always effectively reduces the integral over a finite (often not contiguous) area of the *uv* plane. In particular, the sampled *uv* plane is restricted to a minimum *uv* distance that cannot be shorter than the antenna size and a maximum *uv* distance given by equation 5.2.

5.3 Challenges in 21 cm observations

5.4 Array design

Bibliography

[1] A. Richard Thompson, James M. Moran, and Jr. Swenson, George W. *Interferometry and Synthesis in Radio Astronomy, 3rd Edition*. 2017.

Chapter 6

Chapter title

Author Name

Abstract

This chapter discusses some important things

6.1 A Section

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis eu egestas erat. Maecenas tincidunt lacinia tincidunt. Mauris id lectus nec neque feugiat condimentum vitae at diam. In vel orci nunc, non commodo mauris. Vivamus ipsum enim, vulputate quis pharetra non, molestie quis felis. Vivamus porttitor placerat turpis at accumsan. Nunc tortor velit, faucibus a rhoncus nec, blandit non elit. Nam consectetur lectus eu nisi blandit dapibus rhoncus dui tempus. Mauris fermentum dolor vel ipsum vulputate sit amet ultricies tortor lacinia. Donec ut nibh erat. Morbi nec mi ante. Integer nec vestibulum diam. Donec tincidunt pellentesque quam, ut interdum mauris venenatis condimentum. Nam condimentum, augue in aliquet gravida, neque dui elementum eros, id semper eros purus sed felis. Curabitur in justo sit amet sapien ultrices hendrerit at quis nibh. Quisque iaculis pulvinar tincidunt.

$$
\begin{aligned}
C(12) &= \left[\vec{\pi} \cdot \vec{\phi}\,(x+r) \right] \\
&\approx 1 - \text{const} \frac{r^2}{L^2} \int_r^L \frac{x\mathrm{d}x}{x^2} + \cdots \\
&\approx 1 - \text{const} \frac{r^2}{L^2} \ln \frac{x\mathrm{d}x}{x^2} + \cdots .
\end{aligned}
\tag{6.1}
$$

Aenean tellus risus, porta sit amet porta vitae, tincidunt ut felis. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Phasellus pulvinar placerat velit auctor egestas. Vivamus euismod fringilla tincidunt. Sed ut magna felis, id sollicitudin nunc. Quisque a dui eu erat consectetur egestas a quis justo. Aenean euismod congue diam, vel posuere urna fermentum sit amet. Lorem ipsum dolor sit amet, consectetur adipiscing

Figure 6.1: This is figure 1 in chapter 1.

Table 6.1: Greek Letters.

α	β	γ	δ	ε	ϵ	ζ	η
θ	ϑ	γ	κ	λ	μ	ν	ξ
o	π	ϖ	ρ	ρ	σ	ς	
τ	υ	ϕ	φ	χ	ψ	ω	
Γ	Δ	Θ	Λ	Ξ	Π	Σ	Υ
Φ	Ψ	Ω					

elit. Mauris faucibus lacus eget est mollis auctor. Donec at nibh ligula, et posuere massa. Phasellus quis leo diam [1]. Donec aliquam blandit risus, eu venenatis ante euismod eu. Curabitur cursus justo id arcu condimentum feugiat. Integer sapien urna, vulputate et adipiscing nec, convallis et justo. Suspendisse in ipsum at felis ornare interdum [2],

Cras adipiscing sagittis nunc vel luctus. Suspendisse volutpat augue quis erat semper consequat dignissim tellus euismod. Morbi hendrerit, tellus id aliquam iaculis, nibh leo tincidunt eros, vitae varius ligula felis in mi.
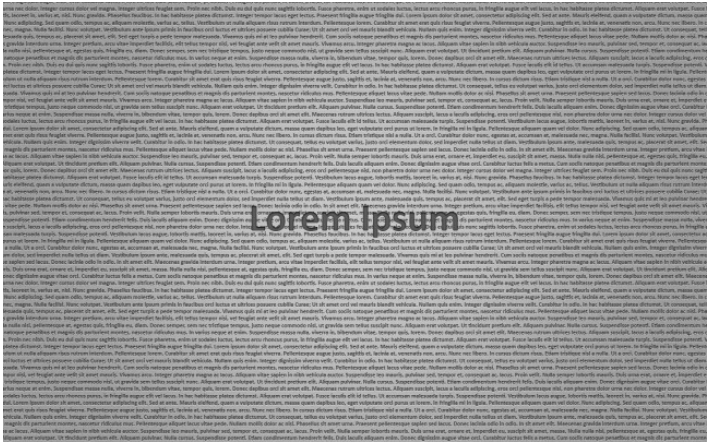
Figure 6.2: This is figure 2 in chapter 1.

Bibliography

[1] KI Diamantaras and SY Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc. New York, NY, USA, 1996.

[2] D. Tulone and S. Madden. PAQ: Time Series Forecasting for Approximate Query Answering in Sensor Networks. In *Proceedings of the 3rd European Workshop on Wireless Sensor Networks*, pages 21–37. Springer, 2006.

Chapter 7

Chapter title

Author Name

Abstract

This chapter discusses some important things

7.1 A Section

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis eu egestas erat. Maecenas tincidunt lacinia tincidunt. Mauris id lectus nec neque feugiat condimentum vitae at diam. In vel orci nunc, non commodo mauris. Vivamus ipsum enim, vulputate quis pharetra non, molestie quis felis. Vivamus porttitor placerat turpis at accumsan. Nunc tortor velit, faucibus a rhoncus nec, blandit non elit. Nam consectetur lectus eu nisi blandit dapibus rhoncus dui tempus. Mauris fermentum dolor vel ipsum vulputate sit amet ultricies tortor lacinia. Donec ut nibh erat. Morbi nec mi ante. Integer nec vestibulum diam. Donec tincidunt pellentesque quam, ut interdum mauris venenatis condimentum. Nam condimentum, augue in aliquet gravida, neque dui elementum eros, id semper eros purus sed felis. Curabitur in justo sit amet sapien ultrices hendrerit at quis nibh. Quisque iaculis pulvinar tincidunt.

$$\begin{aligned}
C(12) &= \left[\vec{\pi} \cdot \vec{\phi}\,(x+r) \right] \\
&\approx 1 - \text{const}\frac{r^2}{L^2} \int_r^L \frac{x\,\mathrm{d}x}{x^2} + \cdots \\
&\approx 1 - \text{const}\frac{r^2}{L^2} \ln \frac{x\,\mathrm{d}x}{x^2} + \cdots .
\end{aligned} \tag{7.1}$$

Aenean tellus risus, porta sit amet porta vitae, tincidunt ut felis. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Phasellus pulvinar placerat velit auctor egestas. Vivamus euismod fringilla tincidunt. Sed ut magna felis, id sollicitudin nunc. Quisque a dui eu erat consectetur egestas a quis justo. Aenean euismod congue diam, vel posuere urna fermentum sit amet. Lorem ipsum dolor sit amet, consectetur adipiscing

Figure 7.1: This is figure 1 in chapter 1.

Table 7.1: Greek Letters.

α	β	γ	δ	ε	ϵ	ζ	η
θ	ϑ	γ	κ	λ	μ	ν	ξ
o	π	ϖ	ρ	ρ	σ	ς	
τ	υ	ϕ	φ	χ	ψ	ω	
Γ	Δ	Θ	Λ	Ξ	Π	Σ	Υ
Φ	Ψ	Ω					

elit. Mauris faucibus lacus eget est mollis auctor. Donec at nibh ligula, et posuere massa. Phasellus quis leo diam [1]. Donec aliquam blandit risus, eu venenatis ante euismod eu. Curabitur cursus justo id arcu condimentum feugiat. Integer sapien urna, vulputate et adipiscing nec, convallis et justo. Suspendisse in ipsum at felis ornare interdum [2],

Cras adipiscing sagittis nunc vel luctus. Suspendisse volutpat augue quis erat semper consequat dignissim tellus euismod. Morbi hendrerit, tellus id aliquam iaculis, nibh leo tincidunt eros, vitae varius ligula felis in mi.

Figure 7.2: This is figure 2 in chapter 1.

Bibliography

[1] KI Diamantaras and SY Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc. New York, NY, USA, 1996.

[2] D. Tulone and S. Madden. PAQ: Time Series Forecasting for Approximate Query Answering in Sensor Networks. In *Proceedings of the 3rd European Workshop on Wireless Sensor Networks*, pages 21–37. Springer, 2006.

Chapter 8

Status of 21 cm interferometric experiments

Cathryn M. Trott (ICRAR-Curtin), Jonathan Pober (Brown University)

Abstract

Interferometric experiments of the reionization era offer the advantages of measuring power in spatial modes with increased sensitivity afforded by multiple independent sky measurements. Here we review early work to measure this signal, current experiments, and future opportunities, highlighting the lessons learned along the way that have shaped the research field and experimental design. In particular, this chapter discusses the history, progress, challenges and forecasts for detection and exploration of the spatial structure of the 21 cm brightness temperature signal in the Epoch of Reionisation using interferometric experiments. We discuss GMRT, PAPER, LOFAR, MWA, and the future HERA and SKA.

8.1 Introduction

Because they provide both rapid mapping speed and good angular resolution, interferometers have become the preferred instrument for experiments looking to measure the expected spatial fluctuations in the 21 cm signal. The current instruments hosting such experiments include the Murchison Widefield Array, MWA[1] [17, 115, 52]; the Precision Array for Probing the Epoch of Reionization, PAPER[2] [94]; the LOw Frequency ARray, LOFAR[3] [124, 100]; and the Long Wavelength Array, LWA[4] [36]. In the future, we expect the Hydrogen Epoch Reionization Array, HERA [31] and the Square Kilometre Array, SKA-Low [65]. Sensitivity predictions for most of the current experiments (e.g. [11, 103]) find that they will not be capable of achieving the necessary signal-to-noise to image the 21 cm signal directly (although

---

[1]http://www.mwatelescope.org

[2]http://eor.berkeley.edu

[3]http://www.lofar.org

[4]http://lwa.unm.edu

see [134] for a study with LOFAR). As such, most of these experiments are targeting a detection of the 21 cm power spectrum, which can be constrained with higher signal-to-noise compared with an image because the isotropy and homogeneity of the Universe allows the 3D k-space power spectrum to be averaged over spherical shells of constant $|k|$. Even using the power spectrum, typical predictions for the requisite observing time are of order 1,000 hours (see Figure 8.1).

However, beyond the need to achieve the requisite sensitivity, experiments are faced with the daunting task of isolating the 21 cm signal from foregrounds that can be up to five orders of magnitude brighter. While the two can, in principle, be separated by their distinct spectral behavior, the inherently frequency-dependent response of radio interferometers complicates the picture significantly. In this chapter, we review the challenges faced by current interferometric 21 cm experiments as well as the progress to-date in overcoming them. The detailed structure of this chapter is as follows. In §8.2, we present the history of experiments and techniques that led to the design of current 21 cm experiments. In §8.3, we discuss the distinct approaches each experiment has developed to overcome the challenges associated with these observations, and in §8.4 we review the current published upper limits on the 21 cm signal strength from these experiments. In §8.5, we highlight the currently unsolved problems at the forefront of experimental 21 cm cosmology and conclude in §8.6 with a discussion of the potential for both current and future experiments to overcome them.

8.2   Early work

The origins of the approaches that current experiments are taking to detect the Epoch of Reionization power spectrum can be traced to the development of radio interferometry observational techniques and Cosmic Microwave Background (CMB) analysis methodology.

Radio interferometers measure the cross-correlation of voltages detected with two antennas, extracting the sky signal in a complex-valued dataset that encodes sky emission location and intensity, and as a function of antenna separation vector and frequency [113]. For small field-of-view instruments (large antenna aperture), the measured signal is well-approximated as the 2D Fourier Transform of the sky brightness, attenuated by the antenna response function (the primary beam).

Motivated by analysis of CMB datasets in the 1990s and 2000s, and the curved nature of full-sky imaging, early discussion of power spectrum estimators used spherical harmonic basis functions to describe the signal and extract optimal estimators [112]. CMB studies suffer from some of the challenges faced also by EoR experiments: wide fields-of-view, low sensitivity, limited angular resolution, and foreground contamination. Unlike EoR, which is an evolving signal in redshift space, CMB studies are single frequency experiments focussed on angular statistics. As such, the foreground mitigation and treatment approaches of CMB studies are of limited use for EoR studies, which attempt to separate foregrounds from the 21 cm signal using the frequency axis. Nonetheless, the fundamental need to extract a weak signal from complex and highly-contaminated data is shared between the two fields, and Tegmark used this experience to apply CMB analysis techniques to early EoR methodology development. Since an interferometer natively measures in Fourier space, there was a transition from the natural basis of curved sky functions (spherical harmonics) to the inter-

ferometer measurement space (Fourier modes) in discussion of optimal estimators for EoR science [73].

This work was supported by groundwork laid out for doing EoR power spectra with radio interferometers, including cosmological and unit conversions [81, 94] and noise considerations for astrophysical parameter estimation with specific future experiments [76]. McQuinn and colleagues discussed a simple foreground model where fitting of a smooth spectral function could remove their effect cleanly, focussing on array sensitivity as the limiting factor for future experiments. However, the lack of any real-world experiments attempting the detection meant they failed to realise the extent of foreground spectral contamination.

More sophisticated approaches to foreground modelling and mitigation appeared in the mid-2000s, with [19] beginning a set of papers that explored the signature of smooth spectrum sources in the EoR power spectrum parameter space. Initially, low-order polynomials were explored to fit and remove these sources. However, lacking a physical motivation for this functional form to robustly separate foregrounds from cosmological signal, polynomials were replaced with more realistic functions. Ultimately, the likelihood of removing not only foregrounds but also cosmological signal when fitting and subtracting models, particularly considering the large difference in magnitude of the two signals, has steered the research field away from this approach to foreground mitigation.

As part of this better appreciation for the impact of foregrounds, particularly with the knowledge that they are used also for data calibration, [29] explored the required accuracy of source models such that foregrounds may be subtracted to a level sufficient to detect the EoR. This work was the first to show the characteristic wedge in power spectrum parameter space, a triangular region in angular and line-of-sight wavenumber space representing the signature of smooth-spectrum sources observed with an interferometer.

8.3 Experimental methodologies and current experiments

In this section we introduce the different instruments that have previously taken, or currently are taking and analysing, data for an interferometric EoR experiment. We start by presenting the relevant parameters of the telescopes that these experiments use, highlighting and motivating the different observational and analysis approaches taken by each. Table 8.3 lists the location (including latitude), frequency (redshift) range, number of stations/antennas, station diameters, and maximum baseline, and field-of-view at 150 MHz for the relevant instruments. Italicised telescopes are discussed in this Chapter. We also plot the full uncertainties (including sample variance) for a 1000 hour observation at z=8.5 (10 MHz bandwidth) for each experiment as a function of spatial wavenumber in Figure 8.1. We uniformly assume that the modes within the horizon are inaccessible due to foreground contamination, and note that this is a broad assumption that is not applicable to all experiments. Note also that MWA's and HERA's large fields-of-view gives them access to smaller wavenumbers. This figure also includes a nominal signal strength (black, 21cmFAST, [78]), but this level is highly uncertain. The proximity of the curves to this line highlights the difficulty with predicting the real sensitivity of experiments, particularly in light of the large number of observing hours required to reach an expected detection. The parameters shown in the table, and the curves shown in Figure 8.1 motivate and frame the discussion of different experiments in the follow-

Facility	Location (Latitude)	Freq. [MHz] (z)	N_{ant}	Max. baseline	FOV_{150}
GMRT	India (19.1^oN)	150–300 (3.7–8.5)	30	30 km	2.5^o
MWA	Australia (26.5^oS)	70–90, 135–195 (15–19, 6–10)	128	5 km	25°
LOFAR	Netherlands (52.9^oN)	30–80, 120–190 (17–46, 6–11)	50–60	50 km	5°
PAPER	South Africa (30.6^oS)	110–180 (7–12)	32–64	210 m	60°
LEDA[1]	USA (34^oN)	45–88 (15–30)	256+	<10 km	70°
21CMA	China (42^oN)	50–200 (6–27)	81	6 km	10°

Table 8.1: General parameters for the telescopes undertaking interferometric observations of the Cosmic Dawn and EoR. Italicised telescopes are discussed in this Chapter. [1]LEDA is a total power experiment using interferometry for data calibration.

ing sections. Experiments are forced to undertake different approaches to observations and data analysis, because the physical limitations of the systems promote different systematic errors into the forefront for each experiment. There is no silver bullet telescope for undertaking this experiment, however, and after reviewing the main experiments, we discuss the pros and cons of different features.

8.3.1   Giant Metrewave Radio Telescope - GMRT

The GMRT [111] is a Y-shaped array of 30 45 m dishes spread over 25 km in western India. Operating between 50 MHz and 1420 MHz, its 153 MHz receiver has been most used for Reionization studies. Motivated by early work in the post-reionization era (325 MHz and 610 MHz receivers) to statistically detect 21 cm fluctuations and understand the foreground contamination to these data, the low frequency receiver opens the door to exploring the Reionization era. The methodology developed has focused on angular power spectra, measured at a range of frequencies, and pioneered much of the early work to use spectral correlation of foregrounds as a way of treating them. With a lack of short baselines and poor instantaneous *uv*-coverage (Figure 8.2), the array is suited to building high resolution foreground models, and computing the foreground angular correlation function ([107]).

GMRT work has largely utilised the visibility correlation function, which cross-correlates visibilities to study the spectral and spatial structure of the sky. Visibility correlation functions were also explored for 21CMA analysis [136]. Unlike other experiments, which have cross-correlated interleaved time samples to remove noise power bias, GMRT has usually opted for cross-correlating visibilities from adjacent frequency channels. This has different systematics, with finer spectral resolution required to minimise visibility decorrelation.

During the 2000s, there was a series of papers developing a formalism for use of this visibility correlation function to measure angular modes. [14] introduced a cross-visibility angular correlation function to measure HI fluctuations post-reionization. [15] then related the cross-visibility correlation function across baselines and frequencies to the power spectrum of brightness temperature fluctuations, presenting the full formalism and expected results in different epochs. They suggest that the cosmological is uncorrelated for frequency channel differences larger than 1 MHz, allowing signal to be 'easily distinguished from the continuum sources of contamination'. [4] then extended this formalism to model the expected
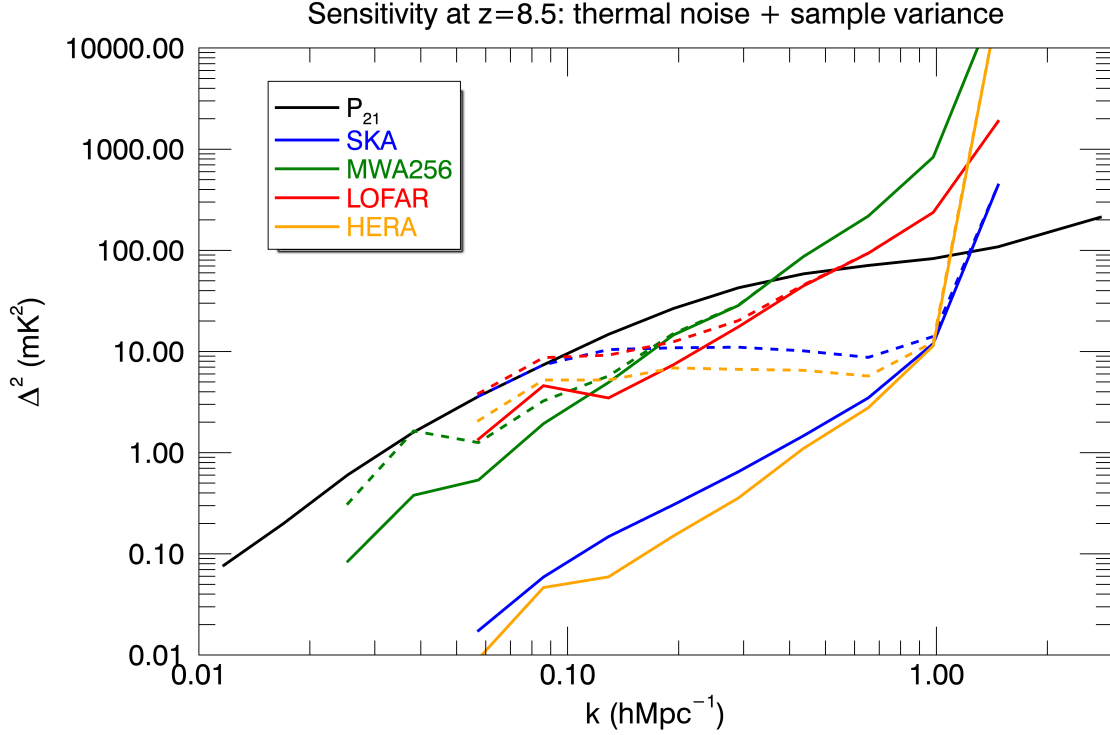
Figure 8.1: Estimated total uncertainty on the dimensionless power spectrum for a 1000 hour observation with several experiments at z=8.5 compared with a model input power spectrum (21cmFAST,[78]), and assuming that modes within the horizon are inaccessible. Solid lines are only thermal noise uncertainty, while dashed lines include sample variance. (Black) Model cosmological signal; (blue) SKA; (green) MWA256; (red) LOFAR; (orange) HERA.

foreground continuum signatures in the cross-correlation visibility space, and compared with GMRT observations. Their results were hindered by calibration errors, which caused decorrelation of the signal over frequency, but presented the first application of this technique to data. [30] provided an extension of the visibility cross-correlation approach to estimating power spectra to a multi-frequency angular power spectrum (MAPS), utilising decorrelation of signals over frequency to extract information about bubble sizes and distributions as a function of redshift while suppressing the effects of foreground contamination.

[41] published the first measurement of post-reionization neutral hydrogen fluctuations with GMRT (HI intensity mapping) at z=1.32 (610 MHz) and using the MAPS formalism. They used a fourth-order polynomial to remove smooth foregrounds, in line with early attempts with many experiments to fit a parametric function without physical motivation. [42] then demonstrated improved foreground removal for 610 MHz observations by tapering the primary beam function and reducing sidelobes; [44] further extended the work to the reionization epoch using 150 MHz observations to characterize the foregrounds with the MAPS formalism.

Using an alternative analysis to the MAPS formalism, [92] analysed 50h of data at

z=8.6 with a simple piecewise linear foreground subtraction method and cross-correlation of foreground-subtracted images. The result was a reported upper limit on the 21 cm signal strength of $(70 \text{ mK})^2$. However, [91] re-analysed the data with a more sophisticated foreground subtraction technique, including a calculation of signal loss due to foreground fitting. The result was an increase in the upper limit to $(248 \text{ mK})^2$, indicative of the degree to which signal loss can affect results.

More recently, [26] published a series of papers introducing and exploring the use of two new optimised power spectrum estimators using visibility correlations: the Tapered Gridded Estimator (TGE) and Bare Estimator (BE). The key concept for the TGE, which has been further discussed in the literature in subsequent papers [27], is to use a Fourier beam gridding kernel that is larger than the physical beam kernel, thereby decorrelating sources at the edge of the field-of-view. Note that this approach is not a silver bullet to removing the effect of horizon sources, because their sidelobes remain in the data even if they have been attenuated. Originally developed as angular power spectra as a function of frequency, the TGE work has recently been expanded to use the line-of-sight spatial information [16]. The BE directly squares adjacent visibilities to provide individual measurements of the power, but this has not been used further, possibly due to the large number of visibilities that are accumulated and stored.

Additionally to power spectra, [108] predicted the amplitude of a bispectrum signal with GMRT using its shortest baselines by modelling non-linear clustering. They predicted the signal strength to be comparable to the power spectrum and detectable in 100 hours but this project has not been explored observationally with this instrument.

8.3.2   Murchison Widefield Array - MWA

The Murchison Widefield Array (MWA) is a 256-element interferometer in the Western Australian desert. In Phase I of the array, operating from 2013–2016, it was composed of 128 tiles of 16 dual-polarization dipoles, spread over 3 km [115]. Phase II (2016–) expanded the array to 256 tiles, with longer baselines for improved survey science and sky model building (5 km), and two hexagonal sub-arrays of 36 tiles with short spacings available for redundant calibration and improved EoR sensitivity [126]. It operates in two distinct modes: Extended Array (128 tiles with long baselines), and Compact Array (128 tiles with short baselines including two 36-tile redundant subarrays in a hexagonal configuration). The Compact Array is principally used for EoR science (see Figure 8.3). The MWA is a general science telescope, with multiple science goals [17]. As such, it balances high surface brightness sensitivity on EoR scales, redundant and non-redundant elements, and longer baselines for good imaging capabilities. The instantaneous *uv*-coverage of the MWA is excellent, allowing for science-quality snapshot imaging (2-minute). The MWA is also a wide-field instrument, with a field-of-view of 25 degrees at 150 MHz. This wide field-of-view, combined with the complex frequency-dependent shape of the aperture array primary beam, and analogue electronics, create challenges for data analysis. The two-stage analogue beamformer produces a frequency bandpass that contains 24 coarse channels over a 30.72 MHz bandwidth (chosen from the full bandwidth listed in Table 8.3), with missing regular channels between the coarse bands. This instrumental spectral structure provides a challenge to producing instrumentally-clean output EoR datasets.
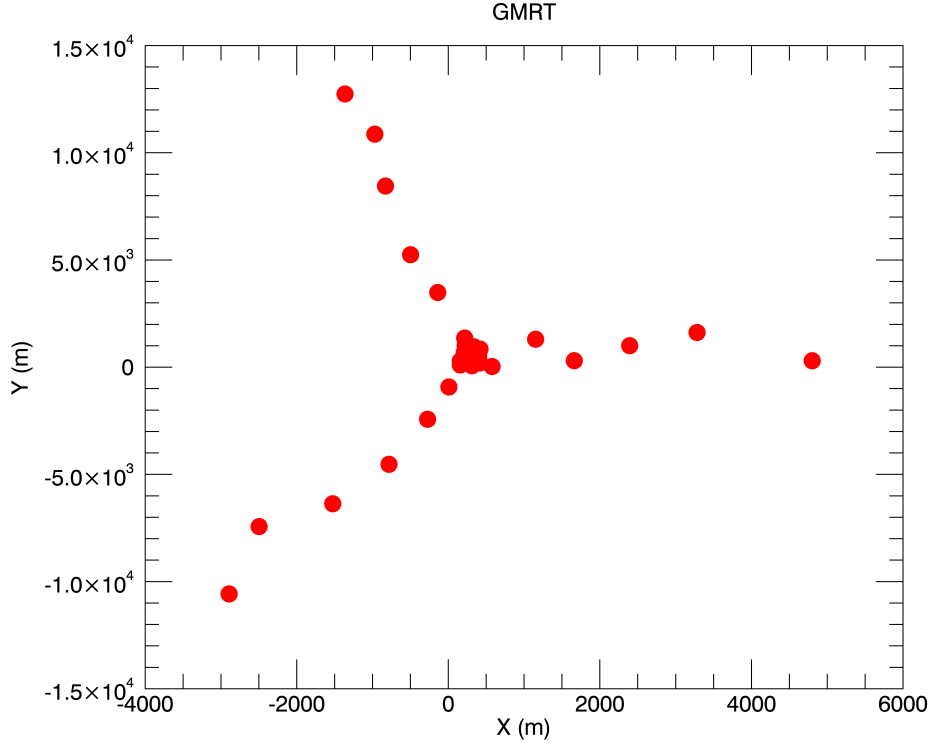
Figure 8.2: Array configuration for the GMRT: thirty 45 m dishes spread over 30 km.

Early deployments of the array, with 32 tiles, were used for preliminary science, and to begin to survey the EoR fields [128]. Upon completion of the 128 tiles, the MWA Commissioning Survey provided the first sky catalogue for use for calibration of EoR data [49]. This work paved the way for the GLEAM survey [125] and catalogue [48], yielding 300,000 sources in the southern sky. GLEAM provides the basis for the current sky model for point sources in EoR observations, augmented by individual models for extended sources.

In line with developments in concurrent experiments, prior to data acquisition the MWA EoR collaboration focused on relatively simplistic foreground fitting and removal methods, but with an increasing understanding of the signature of smooth-spectrum foregrounds in the wavenumber parameter space of an interferometer [19, 29, 118]. There are now two primary EoR data calibration and source subtraction pipelines used by the collaboration: the Real-Time System (RTS, [79]) and Fast Holographic Deconvolution (FHD, [110]). Both use underlying catalogues of sources that have been generated by cross-matching multiple low-frequency sky catalogues. PUMA [72] generates an observation-specific sky model of point sources and double sources [106], and includes shapelet-based and point source-based models for extended sources. The RTS calibrates the data in two steps, both of which rely on a weighted least-squares minimisation: (1) overall direction-independent (flux density and phase) calibration on a full model of 5,000 sources; (2) direction-dependent corrections along the line-of-sight to bright sources. The direction dependent corrections are then ap-
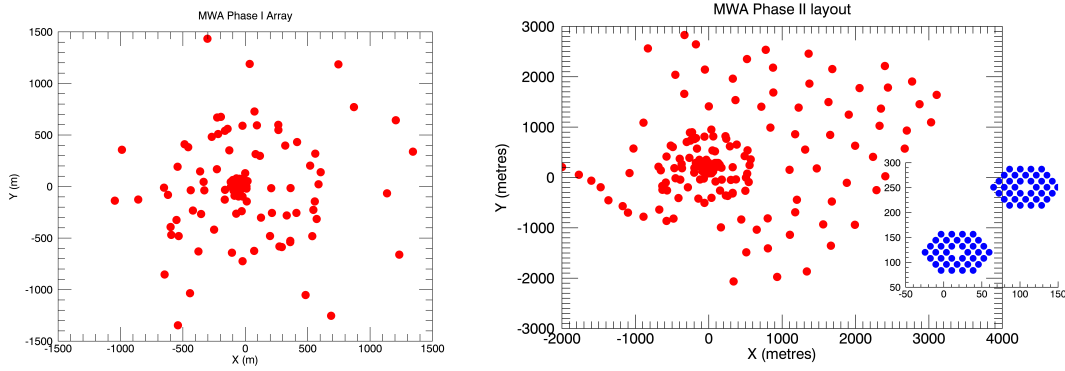
Figure 8.3: Array configurations for the MWA (Phase I, left; Phase II, right, including cutout of hexagonal subarrays (blue)): 128 (256) 4.4 m aperture array tiles spread over 3 (5) km.

plied to sources in the region of the fit, and the 5,000 source sky model is subtracted. FHD calibration [110] computationally optimizes the direction-dependent and wide-field imaging steps by pre-computing the mapping from Fourier to real space. FHD relies on an underlying point source sky model [21], which generates an observation-specific calibration model for >10,000 sources based on the GLEAM catalogue and other cross-matched surveys.

Early developments of power spectrum pipelines stemmed from the inverse covariance quadratic estimator framework pioneered in CMB studies [112], and applied to theoretical EoR datasets by Liu & Tegmark in 2011 [73]. This work was further developed by Dillon in a series of papers that explored how to bridge some of the differences between the ideal estimator and a physical dataset [34, 33]. In particular, Dillon discussed missing data, and large data volumes. An adapted approach was then applied to three hours of MWA data, showing promising results [33].

One key feature of the optimal quadratic estimator formalism is the whitening of data according to the correlated covariance introduced by the uncertainty on residual foregrounds. This is effectively a down-weighting of data that are heavily affected by foregrounds, thereby improving signal-to-error. Subsequent analysis of a higher redshift dataset was used to estimate the principal eigenmodes of the data in spectral space, identifying these with bright foregrounds [34]. The covariances of these modes were then used in the estimator to down-weight and decorrelate data, yielding improved limits at $z = 6.8$. However, as with commensurate and subsequent work with PAPER that used this technique, it had the large potential to cause bias in the estimates. Re-use of the same dataset to empirically estimate the data covariance, and then fit for it, causes re-substitution bias, a well-known statistical effect where the performance of an estimator can appear much better than it actually is. In this work, Dillon was careful to estimate covariances empirically while omitting the *uv* cells in question, to avoid bias, however there was still limited information available in the remaining cells. Thus, although this work was careful to not try to subtract the foreground bias directly, use of the empirical covariance in the data weighting, and lack of a full end-to-end simulation to demonstrate no signal loss, makes this approach prone to large bias. It has not been used to analyze MWA data since the original analysis in [34].

In a later paper, describing the CHIPS estimator, Trott also developed an inverse co-

variance quadratic estimator formalism using a model foreground covariance [116]. Unlike the empirical approach of earlier work, this does not use the data itself to form the foreground covariance, but a model for the expected spatial and spectral structure of point source foregrounds. However this approach can suffer from similar effects, whereby error in the covariance can propagate into the analysis. Therefore, this inverse foreground covariance has never been applied to data used in publication due to the output's sensitivity to the choice of foreground model.

A second principal power spectrum estimator for MWA EoR analysis, εppsilon, was independently developed from CHIPS [9]. εppsilon prioritizes the propagation of thermal noise error from the visibilities (with estimates provided by FHD) through to the power spectrum while also providing a suite of diagnostics for assessing the performance of the estimator in a number of domains.

Both εppsilon and CHIPS (without the foreground covariance weighting) were used in the EoR limit paper led by Beardsley [12], which processed 32 hours of MWA Phase I high-band data to power spectrum limits. At the time, these results were highly-competitive in the field, but the data were clearly still systematic-dominated. At a similar time, Ewall-Wice published the first measurement of upper limit from the Cosmic Dawn (Epoch of X-ray heating, EoX) from 3-hours of MWA data above $z = 15$ [37].

One of the clear outcomes of the early upper limit publications from MWA (and other instruments, particularly LOFAR) was that the data were highly systematic-dominated in modes relevant for EoR, and accumulating more data into the power spectrum estimator would offer no advantage. With this realisation, the MWA collaboration embarked on a two year program to prioritize understanding and treating systematics over processing large datasets, despite more than a thousand hours having been collected by the instrument. This work encompassed (1) improving the sky model (point, extended and multiple sources, [106, 123]); (2) understanding the impact of calibration choices on residuals and uncertainties [10, 117, 119, 38, 84]; (3) improving the primary beam modelling [71]; (4) developing data quality metrics for data triaging (RFI, ionospheric activity, [59, 121, 127]), (5) developing and refining redundant and hybrid calibration pipelines for Phase II [70, 60, 20]. A final important step was the development of a full end-to-end simulation to demonstrate that there was no signal loss in the chain from telescope to data product. The results of this work include upcoming EoR limits from re-analysis of Phase I data and new Phase II data, as well as exploration of new techniques for exploring the EoR [122].

A final, key insight from recent work helps to address the current questions in the EoR research field about robustness of any future claimed detection of cosmological signal. Along with confirmation by other telescopes, ability to detect the same signal in independent observing fields, where the foregrounds are different, is crucial. In [120], MWA data from two observing fields was studied with a Kernel Density Estimator to understand the similarities and differences between the statistical structure of data from independent sky areas. This work can lead to a better understanding of robustly discriminating contamination from cosmological signal.

8.3.3 Low Frequency Array - LOFAR

LOFAR is a composite aperture array low-frequency radio interferometer. It has two primary station types; the High-Band Antennas (HBA, 120–190 MHz) and Low-Band Antennas (LBA, 30–90 MHz). Both station types have been used for EoR and Cosmic Dawn science. Athough LOFAR formally contains baselines of thousands of kilometres to the international stations, it is only the Dutch-based stations that are used for actual EoR measurements. Figure 8.4 shows the central stations (blue cut out) and the nearest remote stations (red).
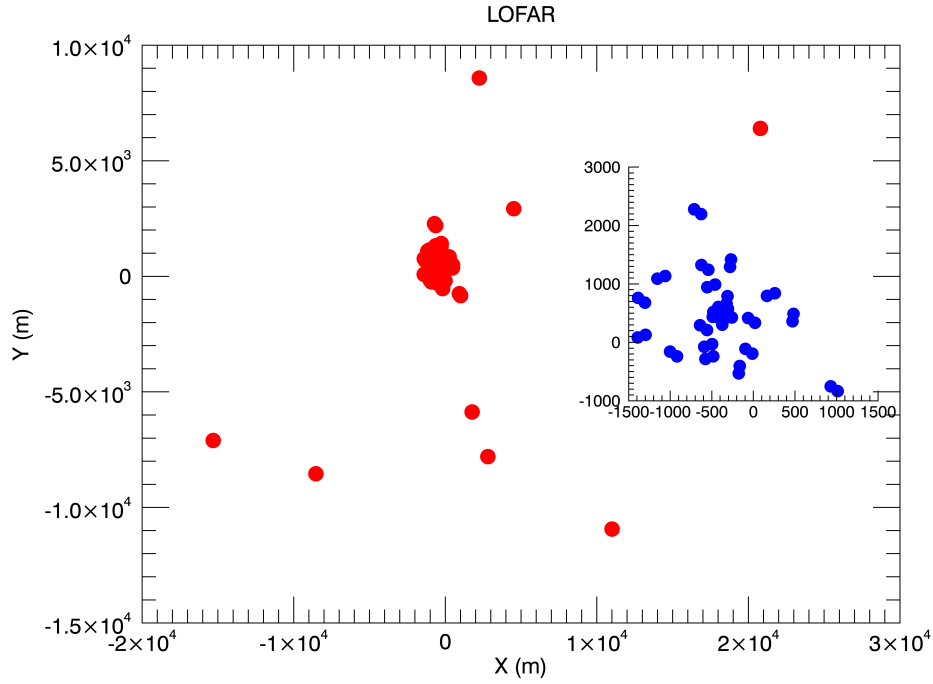

Figure 8.4: Array configuration for the central and inner remote stations of LOFAR (red), and subplot showing the central stations only (blue): 30–40 m aperture array dipole stations spread over tens of kilometres.

The LOFAR latitude allows for circumpolar observations with long winter nights. As such, one of the primary observing fields is the North Celestial Pole, which can be observed for more than 12 hours in the winter months. Early work with the LOFAR EoR Key Science Project focussed on foreground mitigation, choice of observing fields, and data analysis methodology. As with many of the published papers in this early epoch, foregrounds in [57] were modelled to be subtracted with a simple smooth fitting function. This work is notable because it provided realistic models for a range of different foreground components, and included discussion of the treatment of polarized foregrounds.

[58] extended the work from 2008 to focus on simulations of Faraday Rotation from polarized Galactic foregrounds. FR rotates the phase of the intrinsically-smooth foreground component yielding spectral structure that may mimic the EoR signal. In this work, Jelić

shows the effect of inaccurate data calibration on polarized emission, which can imprint total intensity structure if the polarized instrumental response is incorrect.

The SAGE algorithm (Space Alternating Generalized Expectation Maximization) was first introduced in 2011 by [61], and provides the basis for all calibration of LOFAR EoR datasets to the present day. Based on the well-known Expectation Maximization (EM) algorithm, which iteratively fits for calibration parameters (maximizes the likelihood with respect to a set of parameters and then alters the parameters to find a new likelihood) when the underlying system model contains unobserved variables, SAGE extends the traditional least-squares fitting to allow for more model flexibility, and improved convergence and efficiency. Part of the SAGE algorithm performs direction-dependent calibration towards clusters of sources on the sky, thereby allowing for ionospheric distortion of the sky model.

Detailed total intensity and polarized imaging of the LOFAR EoR fields were presented in [130] and [56]. The North Celestial Pole field allows for deep, long winter nighttime observations, and was shown to be able to be calibrated over long integrations. Observations of low Faraday depth structures in the ELAIS-N1 field yielded structures that would be problematic for EoR science if the degree of polarization leakage into total intensity exceeded 1%. This quantification of the accuracy required of instrumental polarization models was the first of a set of papers that explored polarized signal and leakage for EoR science. Thus far, the LOFAR EoR collaboration has undertaken the most extensive work to quantify the impact of polarization leakage, while the MWA collaboration has made some observations of polarized emission in their data ([69], [68], [13]). In [8] and [7], Asad and colleagues first studied the polarized emission in the 3C196 EoR field, finding them to be localized around a small Faraday depth, and quantified the leakage, and then studied the accuracy of the LOFAR polarized beam model to be able to limit leakage into total intensity. Given the level of polarized to total intensity, and a beam model accurate to 10% at the field centre, the leakage and subsequent spectral structure would be acceptable for EoR science. Finally, [6] considered the more problematic impact of wide-field polarization leakage on the EoR power spectrum. Far from the field centre, the primary beam models are less accurate, and sources imprint additional spectral structure due to the chromaticity of an interferometer. In these cases, bias was found to persist in the EoR power spectrum.

In early work on fitting foregrounds, Harker and colleagues [46] discussed the use of Wp smoothing as a non-parametric method for fitting a smooth function, based on limiting the number of inflection points in the fit. Further, they discuss the systematic errors introduced by the fitting routine, methods for estimating these, and for accounting for them in the final uncertainties. This approach is used again in the work of [77] for the Gaussian Process Regression fitting, and also is used generically in the PAPER analysis to try to understand signal loss. There, and elsewhere, use of the same dataset to empirically estimate the bias, and then to correct it, leads to resubstitution bias, underestimate of bias and signal loss.

In a series of papers, Chapman and collaborators explored novel approaches to fitting and removing foreground signal from image-based datacubes. In [24], they introduced the FastICA technique, as a non-parametric method that estimates independent foreground components and their mixing for each image pixel and frequency. The advantage of such methods is that they do not rely on any a priori knowledge of the signal, but instead only assume that the full signal can be represented by a small number of components (sparsity), thereby allowing for good estimation with a given dataset. The disadvantage lies in the sensitivity of

results to the number of components the user chooses that the data should contain, and the potential for signal loss if the projection of the estimated components onto the EoR signal is non-negligible. ICA generically minimizes Gaussianity, thereby enforcing smoothness in the fitted components. Beyond ICA, a generalized method (GMCA; Generalized Morphological Component Analysis) was applied ([23, 22]) to use an underlying blind wavelet decomposition of the components, combined with the sparsity and mixing model methodology of the ICA method. As with other methods when the underlying structure, spatial distribution and amplitude of the cosmological signal and foreground components is unknown, the potential for signal loss is present. The ICA and GMCA methods both assume that the cosmological signal has negligible amplitude and is absorbed in the noise. Structural deviations from this assumption can lead to signal loss.

In a new approach to foreground treatment, Mertens and colleagues discuss use of the well-known Gaussian Process Regression (GPR) technique to fit for foregrounds using only an understanding for the spectral data covariance of different components [77]. Unlike parametric methods that assume an underlying model, GPR (an extended version of kriging, which interpolates data based on their known covariance properties) relies only a statistical separation of foreground and cosmological signal via their spectral correlation lengths. This method also suffers from the potential for cosmological signal loss, but the authors attempt capture the potential bias statistically through increased noise. The ultimate utility of this approach has yet to be demonstrated on a large dataset at the time of the writing.

Along with the power spectrum as a measure of the signal variance as a function of spatial scale, the variance statistic was explored with simulations in [102]. The variance of the brightness temperature, as the wavenumber integral over the power spectrum, quantifies the variability in the cosmological signal on the imaging scale (autocorrelation function). Although it provides limited cosmological information, detection of this variance can be theoretically obtained with fewer observing hours. Bayesian power spectrum extraction techniques were also explored in [43], with a view to allowing for a spatially-smooth component to capture the unmodelled diffuse emission in the NCP field. Like other instruments, the data calibration and foreground models were limited to point and extended sources, with the complex diffuse emission difficult to measure and model. Increasingly, the impact of this incomplete sky model has become apparent.

In a landmark paper published by Patil and colleagues in 2016 [100] the source of 'excess noise' and diffuse emission suppression in LOFAR data were studied. Excess noise is the identification of increased noise levels in the data post-calibration compared with expectations of thermal noise and Stokes V measurements. Ultimately, the lack of a diffuse model in the calibration sky model allowed for this signal to be absorbed into the gain calibration solutions, thereby yielding a direction-dependent bias and noise in the residual data. To address this problem, the short baselines containing the majority of the diffuse emission could be excluded, however this leads to increased noise on these scales (due to statistical leverage; effectively this amounts to additional flexibility in the gain solutions on these scales because they are not used in the modelling). This work was undertaken contemporaneously with that of [10] and [117], which both studied the effect of incomplete sky models and spectrally varying bandpass parameters on calibration and residual signal. The combined outcome of these studies is an understanding of the impact of sky model incompleteness, the need to enforce spectral correlation (e.g., regularization as in SAGECal or smooth model fitting) for

calibration parameter fitting, and the approaches to calibration that can mitigate these.

Further exploration of the impact of calibration frameworks and data treatment were then explored in [83] and [88], with a view to having a complete understanding of the end-to-end data processing of LOFAR EoR data on the path to a detection. Unlike in the previous ten years before real observations were undertaken and thermal noise sensitivity was seen to be the major impediment for EoR detection, the field has come to appreciate the crucial roles of unbiased calibration, sky model completeness, and foreground treatment without cosmological signal loss.

The culmination of the lessons learned from statistical leverage and incomplete sky models was applied to two fluctuation upper limit papers published since 2017. In [99], Patil and colleagues presented competitive results from a small set of data ($\sim$10 hours) at $z = [9.6 - 10.6]$, with the best limit of $(59.6 \text{ mK})^2$. This work reported an excess variance, in line with previous discussions, and the use of Stokes V power to remove noise power. At higher redshifts (lower frequencies), Gehlot and colleagues [40] reported upper limits above $z = 20$, with use of the Gaussian Process Regression foreground fitting technique introduced by [77] for EoR science.

Beyond the power spectrum, LOFAR has explored other tracers of the neutral hydrogen temperature field, namely the ability to produce low angular resolution images [134] and the 21 cm Forest [28]. LOFAR like other current instruments, does not have the sensitivity to directly image neutral hydrogen bubbles at the instrumental resolution. However, by lowering the resolution of images (thereby improving the radiometric noise), Zaroubi and colleagues argue that the largest of bubbles may be detectable at low signal-to-noise ratio on the largest of scales late in reionization. The ability to detect the 21 cm Forest (absorption of continuum radio emission along the line-of-sight to high redshift AGN due to intervening neutral gas) remains a challenge and aim of many current interferometers. Ciardi and colleagues showed that LOFAR would have the ability to detect an absorption feature under ideal conditions. Unlike the statistical detection of the power spectrum of temperature fluctuations, the absorption signal amplitude is determined by the astrophysical conditions close to the gas, namely gas kinetic temperature. Cold gas is able to absorb light more readily than heated gas. The failure of this method to date is primarily due to the lack of any known high-redshift radio-loud AGN ($z > 6$). Given the sensitivity of current instruments, a source with flux density exceeding 10 mJy and cold neutral gas would be required. It is likely that the arrival of SKA will provide both the sensitivity and the detection (and confirmation) of high-redshift radio-loud AGN to be able to undertake this experiment. Of the current interferometric experiments, only LOFAR has sufficient sensitivity to be able to attempt this experiment at all.

The utility of extracting higher-order statistics of the 21 cm brightness temperature field were explored in a simulation study of foreground-subtracted image cubes by [47]. Again, the ability to smoothly treat and remove foregrounds placed the burden of detection on pure noise considerations.

8.3.4 Precision Array for Probing the Epoch of Reionization - PAPER

The PAPER experiment was designed as a testbed for developing novel 21 cm cosmology analysis techniques. PAPER antennas were chosen to be small, single dipoles on elevated

ground-screens to enable reconfiguration of the array, and the system used a flexible digital correlator architecture that could scale as the number of antennas grew [96]. The small antenna sizes was also chosen to limit the frequency evolution of the antenna response over the instrument's $110 - 180$ MHz of usable instantaneous bandwidth. The design and results from an initial 8-station deployment of PAPER in Green Bank, WV, USA were described in [94].

In its earlier stages, PAPER deployed its antennas in configurations designed for imaging, including a single-polarization 16-element 300 m diameter ring in Green Bank used for primary beam measurements in [105]. While the Green Bank array was upgraded to a single-polarization, 32-element array, all subsequent publications came using arrays deployed at the SKA-SA site in the Karoo, South Africa. Highlights of early PAPER studies include the creation of a 145 MHz Southern hemisphere sky-catalog using a single-polarization, 32-element array [53] and a study of the radio galaxy Centaurus A using a single-polarization, 64-element array [109]. In both of these cases, the elements were deployed in a randomized configuration over a circle of 300 m diameter to maximize *uv* coverage.

In 2012, however, members of the PAPER team developed what is now referred to as the "delay spectrum" approach for measuring the 21 cm power spectrum. In the delay spectrum approach, visibility spectra from individual baselines are Fourier transformed and cross-multiplied. [93] demonstrated how these delay spectra can be used as estimates of the 21 cm power spectrum, without ever combining visibilities and making an image. [93] also provided sensitivity estimates for the delay spectrum approach using a 128-element PAPER array. [98] then demonstrated how 21 cm foregrounds isolate into what is now commonly referred to as "the wedge" and included the effects of foreground contamination in the sensitivity study. One consequence of the delay spectrum approach is a higher noise level than alternative approaches: power spectra estimated from individual baselines are averaged together, as opposed to coherently combining all the visibilities and forming a single power spectrum, so noise fluctuations average down more slowly. To make-up for this sensitivity sacrifice, [93] proposed using a "maximum redundancy" configuration, in which antennas are arranged to create multiple copies of the same baseline spacing. These redundant baselines can then be averaged together before squaring, helping the noise level to integrate down faster. Although redundant layouts drastically reduce imaging fidelity, the delay spectrum approach does not requiring imaging and so is, in principle, not affected by this consequence.

The decision was made to reconfigure the PAPER array and test the delay spectrum technique in a maximum redundancy layout. However, a short data set in a single-polarization, 64-element "minimum redundancy" (i.e. random layout) with a 300 m diameter was collected and used to make delay spectra from a range of baseline lengths and orientations in [1]. This analysis demonstrated good isolation of foreground emission to the wedge in 2D cosmological *k*-space, suggesting the promise of the delay spectrum technique.

[95] presented the first deep power spectrum limits from a dual-polarization, 32-element, maximum redundancy array (a grid of 8 columns and 4 rows, with a column spacing of 30 meters and a row spacing of 4 meters). Just over 1000 hours of data were used in the analysis. In addition to the basic delay spectrum formalism, [95] introduced two additional analysis techniques: redundant calibration [2, 74], which was enabled by the redundant layout of the array, and a new technique for removing off-diagonal covariances between redundant baselines. These same techniques were applied to the same data over a range of redshifts in

[55].

The techniques of [95] were then applied to a new, 1000+ hour, dual-polarization, 64-element PAPER data set in [5] (the layout of which is shown in Figure 8.5). This analysis improved upon the redundant calibration technique by using the OMNICAL package [135], replaced the off-diagonal covariance removal technique with an inverse covariance weighting approach using empirically estimated covariance matrices (similar to [34]) and applied a new technique known as "fringe rate filtering" (described in [97]). At the time, the [5] limits on the 21 cm power spectrum were believed to be the most stringent to be published and were followed by two separate publications using their measurement to constrain the temperature of the IGM at $z = 8.4$ [104, 45].

However, re-analysis of the [5] data by [25] revealed a critical error in the analysis: empirically estimated covariance matrices are correlated with the data, and weighting by them can bias the recovered signal low. (As described in §8.3.1, this bias has frequently been referred to as "signal loss" — the idea that an analysis technique can remove 21 cm signal along with foregrounds.) In practice, the signal loss in the PAPER analysis was *very* large (nearly four orders of magnitude of potential EoR signal was suppressed) due to the fringe-rate filtering technique that reduced the number of independent samples used to estimate the covariance matrix. Although the analysis in [5] attempted to estimate signal loss using injection of mock EoR signals into the data, their method missed potential loss caused by data-signal cross terms in the covariance matrix and thus concluded that the original analysis was effectively lossless. Incorrect estimates of both the theoretical and observed noise levels in the data also contributed to the belief that the analysis of [5] was sound.

In light of the analysis in [25], all of the PAPER results in [95, 55, 5] are considered to be invalid and do not place meaningful limits on the 21 cm signal.[1] A re-analysis of the full [5] data set using a lossless analysis is forthcoming, but the limits are not expected to be near the same level as [5]. The PAPER experiment also collected two years of data with a dual-polarization, 128-element array, but due to an increased amount of instrument systematics and failures in the aging system, these data are not expected to be published.

The delay spectrum approach does not allow for high accuracy polarization calibration, which needs to be performed in the image domain. Theoretical studies of the effect of Faraday rotated (i.e. frequency-dependent) polarized emission on the delay spectrum technique were presented in [80] and [87] and studies using PAPER data were performed in [3] and [64]. Overall, the effect of polarized emission on the delay spectrum approach can be quite significant, but the overall amplitude is uncertain as there are few constraints on the polarization properties of the 150 MHz sky at the angular scales probed by PAPER. Ionospheric Faraday rotation can also attenuate the polarized signal in data sets averaged over many nights [75].

---

[1]Although [95] did not use the inverse covariance weighting that was the main source of the problem in [5], its covariance removal technique has not been robustly vetted for signal loss and thus the results are considered suspect at best.
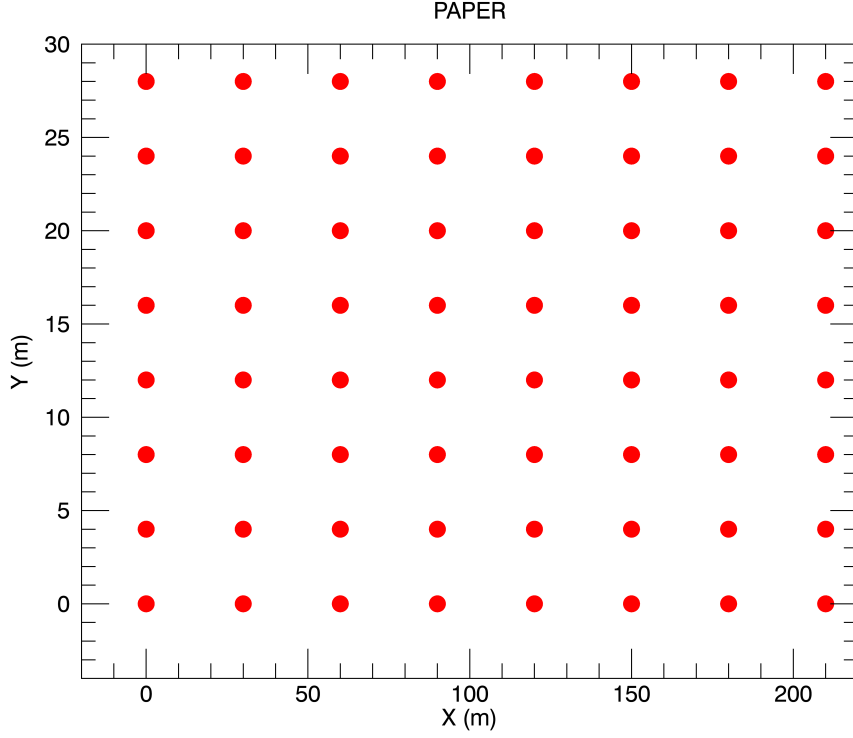
Figure 8.5: Array configuration for the PAPER-64 maximally-redundant array: 64 single dipoles spread over 210 m. Note the distinctly different scales on the *x* and *y* axes.

8.4   Published results

Here we collate the published best limits at each redshift from the current experiments (Table 8.4). PAPER measurements have been omitted. Despite the current published values, there are publications in peer-review now for LOFAR, PAPER, and MWA improving on these results.

8.5   Current challenges

21 cm experiments consist of many components, from the analog telescope design through to power spectrum estimation algorithms. One clear lesson from first generation experiments is that no one aspect of the system can provide the necessary 1-part-in-10^5 dynamic range required to detect the 21 cm signal; rather, the burden needs to be spread across the components of the experiment, alleviating the demands on each of the other components. In this section, we briefly review what we consider five key areas where 21 cm experiments continue to innovate: (1) analog instrument design; (2) data quality control; (3) calibration; (4) foreground mitigation and the associated potential for signal loss; and (5) end-to-end validation of analysis pipelines.

Facility	z	k (hMpc^{-1})	Upper limit (mK)^2	Ref.
MWA	12.2	0.18	2.5×10^7	[37]
MWA	15.35	0.21	8.3×10^8	[37]
MWA	17.05	0.22	2.7×10^8	[37]
MWA	7.1	0.23	2.7×10^4	[12]
MWA	6.8	0.24	3.0×10^4	[12]
MWA	6.5	0.24	3.2×10^4	[12]
MWA	9.5	0.05	6.8×10^4	[34]
LOFAR	10	0.053	6.3×10^3	[100]
LOFAR	9	0.053	7.5×10^3	[100]
LOFAR	8	0.053	1.7×10^4	[100]
LOFAR	23	0.038	6.8×10^4	[40]
GMRT	8.6	0.5	6.2×10^4	[91]

Table 8.2: Best two sigma upper limits on the EoR and Cosmic Dawn power spectrum for each experiment. Only the lowest limits have been reproduced.

1. Analog instrument design. One of the major challenges for 21 cm cosmology experiments is to remove any spectral structure introduced by the instrument that might otherwise mix smooth spectrum foregrounds into the spectral modes occupied by the cosmological signal. One seemingly straightforward approach is to limit the amount of spectral structure in the instrument response through careful analog design. Initial specifications on the HERA system design were to limit spectral structure to a level that would enable the delay spectrum technique without any additional calibration or analysis requirements; however, further study has shown that the HERA design does not meet this stringent specification and will need data analysis algorithms to also model and remove spectral structure from the instrument [31]. The push to larger bandwidths (e.g. $50-250$ MHz for HERA and $50-350$ MHz for the SKA) adds to the challenge of constructing a single instrument with a smooth spectral response over a large range of wavelengths. Analysis with the MWA has also demonstrated how reflections in the analog system can contaminate modes of the EoR power spectrum, suggesting that more stringent specifications on impedance matches and cable lengths are necessary for future instruments [10, 38].

2. Data Quality Control. Given the extreme brightness of human-generated radio signals compared to the 21 cm signal, only a very small number of contaminated measurements are enough to significantly affect the analysis of a large data set. The "gold standard" for identifying radio frequency interference, AOFlagger [90], is used by both LOFAR and the MWA. However, additional quality metrics can still catch corrupted data that slips by this first round of flagging, including ultra-faint, broad-band digital TV transmission [127] and effects due to ionospheric weather [121, 59]. As interferometers grow in size, the large data rates may also require computationally faster algorithms for data quality checks [62]. [89] also demonstrate how even flagged RFI can affect power spectrum analysis if care is not taken.

3. Calibration. Instrument calibration is often regarded as the greatest challenge for existing and future 21 cm experiments. While both the analog design and the methodology

used for power spectrum estimation can ease calibration requirements [82], experiments still need to control the spectral response of their telescopes over wide bandwidths at a level unprecedented in radio astronomy. Typically, antenna-based gain calibration is performed by forward-modeling visibilities and minimizing the difference with the observed data; however, [10], [100], and [117] demonstrate that without additional constraints, calibration performed with an incomplete sky-model can lead to spurious spectral structure in the calibration solutions that can both overwhelm or remove the EoR signal. Redundancy based calibration has been viewed as a promising alternative because it does not reference a sky-model; however, recent work has shown that a sky model is still required to constrain the degeneracies inherent in redundant calibration, and that the same kind of contamination can affect the power spectrum as in sky-based calibration [20, 70, 60]. Calibration of the primary beam response of the instruments is also a major challenge, and several options have been explored, including sky-based calibration [105], using satellite broadcasts [85, 86, 71], and with drones flying transmitters [54].

4. Foreground mitigation. Fundamentally, the real challenges at the heart of 21 cm cosmology come from the intrinsic brightness of the foreground emission. While much of the work to date focuses on removing the instrument response from the foreground spectra, most current experiments use some form of foreground mitigation to help isolate or remove foregrounds. Many distinct approaches have been developed, which can be broadly classified as either "foreground avoidance" and "foreground subtraction." Foreground avoidance methods attempt to isolate foregrounds into the wedge and minimize bleed into the EoR window; power spectra are then only estimated from within the EoR window. Examples of avoidance techniques includes the wide-band iterative deconvolution filter used in PAPER analyses [63] and the inverse covariance weighting techniques also used by PAPER [25]. Foreground subtraction, on the other hand, attempts to remove specific models of the foregrounds — using either real sky catalogs or parametric models for their spectra — while leaving the 21 cm unaffected. Examples of foreground subtraction including the point-source forward modeling and subtraction performed by FHD [9] and the spectral based fitting methods used by LOFAR [22, 77]. It is worth stressing that while these techniques have historically been developed in the context of specific experiments, they are more generally applicable; see [63] for an example of PAPER-developed techniques applied to MWA data and MWA-developed techniques applied to PAPER data.

One of the greatest risks of foreground removal is the inadvertent removal of 21 cm signal, i.e., signal loss. Although many techniques have been developed using frameworks where signal loss is not expected, due to a presumed orthogonality of the foreground description and 21 cm signal basis, there are subtle challenges that arise when faced with a need to achieve five orders of magnitude of dynamic range. While cross-terms between the foreground and signal might have an expectation value of 0, there are still only a finite number of samples going into the analysis, and these cross terms will not have converged to their expectation value — as was the case in the PAPER analysis of [5].

5. Validation. One of the last major challenges for current and future experiments is to rigorously test foreground removal and other analysis algorithms — ideally as part of complete pipeline and not as an independent step — to confirm that 21 cm signal is not being biased or removed. And although foreground removal seems like the step most likely to cause signal loss, it is certainly not the only place that needs further scrutiny. In light of the PAPER

retractions, 21 cm experiments are realizing the importance of simulation-based analysis vetting — ideally with independent, third-party simulations. Many inteferometric simulators exist, including CASA, PRISim [114], OSKAR, and pyuvsim [67]. In turn, it has become important to test the simulators against each other, to verify that they achieve the requisite precision for 21 cm cosmology. These validation efforts can be slow and painstaking, but as experiments push closer to a first detection of the 21 cm signal, they have become more vital than ever. The other avenue for verification is with other instruments and other pipelines providing independent analysis. Use of multiple observing fields can also show robustness to foreground treatment [120].

8.6 Prospects for the future

8.6.1 Current instruments

MWA and LOFAR are both currently pursuing deeper limits. Armed with new calibration and analysis, and critically, a deeper understanding of the effects of different processing approaches, the level of systematics in data are reduced, and more data can be processed to reduce noise. At this stage, it is difficult to predict whether systematics will remain at deeper levels, and if the fundamental limitations of the instrument will preclude a detection. While the reported detection of the Cosmic Dawn global signal from the EDGES experiment [18] suggests that the spatial power spectrum amplitude may be larger than expected, this is highly uncertain, and the flexibility in possible strengths of the signal in the EoR emission part of the spectrum could help or hinder a detection by LOFAR and MWA. Pursuit of the Cosmic Dawn signal from 75–100 MHz observations with the MWA and LOFAR is also underway, but that introduces even greater challenges of large fields-of-view and poor extended source models at those frequencies.

8.6.2 Future instruments

The SKA and HERA offer the future vision for EoR and Cosmic Dawn science. Like LOFAR and MWA, SKA is a general science instrument, being able to produce its own sky model and calibration framework, while needing to balance design with the other science aims of the observatory. HERA, like PAPER, is a custom EoR instrument, being able to design with a complete focus on EoR science, likely requiring external information to provide a full end-to-end calibration and source subtraction element.

The low-frequency telescope of the SKA Observatory, SKA-Low, will be centred at the Murchison Radioastronomy Observatory in Western Australia, on the same radio quiet site as MWA, ASKAP, EDGES and BiGHORNS [65, 32]. Despite being designed for 512 38 m stations (256 dual-polarization dipoles in each station) spread over >40 km, the core region will contain >200 stations within the central 1 km, with exceptional surface brightness sensitivity for EoR and CD science. With a frequency range available down to 50 MHz, the CD will be accessible to $z = 27$, with sub-stations able to be formed to produce the wider fields-of-view and shorter baselines required for early times. With its exceptional imaging capabilities, SKA-Low aims to pursue power spectrum, direct imaging (tomography) and

21 cm Forest studies. The prices to be paid for this highly-capable instrument are the complexity of the data and instrument, and the large data volumes that will be produced from the telescope, and is therefore faces a more severe version of the challenges currently experienced by multi-purpose dipole arrays such as MWA and LOFAR.

HERA [31] is a smaller instrument (although still significantly larger than any of the existing instruments) being constructed in South Africa. It comprises 350 14 m dipole elements spread over <1 km (331 in a 320 m core) for high EoR sensitivity and moderate imaging and calibration needs (19 outriggers). It will primarily pursue the statistical exploration of the EoR and CD using the delay spectrum technique, with some hope for imaging capability and alternate power spectrum analyses.

8.6.3   Future analyses

Although the spatial power spectrum is the primary data product of most current EoR 21 cm experiments, there are other avenues of pursuit to explore this first billion years of the Universe, including an integrated product (the variance statistic, [101]). Direct imaging is beyond the capability of current instruments, demanding a high surface brightness sensitivity and thousands of hours. This will be pursued by the future SKA [65]. However, there are other statistics that can be pursued through the 21 cm line, and also the opportunity for cross-correlating the signal with other tracers of early Universe evolution. The benefit of the latter approach is that the systematic errors may be different between the two tracers, offering an advantage over 21 cm alone.

At early times, the brightness temperature of the 21 cm line, relative to the CMB traces the matter power spectrum, and is highly Gaussian, but at later times the evolution of ionised bubbles dominates the spatial fluctuations and the signal is expected to have non-zero higher order terms [39, 76, 35]. The shape of the temperature distribution function evolves with time and spatial scale, and differs for different underlying models of the evolution of the Universe. As such, probing these non-Gaussian components can provide complementary information to the power spectrum, which, by design, only captures information in the second moment of the distribution [129].

The bispectrum measures the three-point correlation function, and has been shown to encode non-Gaussianity. In early work to study the expected sensitivity of 21 cm experiments to the bispectrum, [133] computed theoretical expectations for a range of instruments, under the assumption of thermal noise only. More sophisticated recent work included the effects of calibration and foregrounds on the ability to detect the signal. In [122], two bispectrum estimators were developed to take a practical approach to estimation with real data, and were applied to 20 hours of data from Phase II of the MWA. This work discussed some of the advantages and challenges of doing such an experiment with real data.

Cross-correlation studies from the early Universe offer the potential for new astrophysical insight and reduced observational biases and errors. In the context of the MWA, [131] used data to explore the cross-correlation of the 21 cm image from the EoR-0 observing field, and the CMB field measured by Planck. An additional tracer that can be used is the population of high-redshift LAEs, which are observable in ionised regions [132, 66, 51]. The SKA's Synergy group is exploring the potential for multi-facility observations, including the exciting prospects available with WFIRST [50].

Bibliography

[1]

[2]

[3]

[4] S. S. Ali, S. Bharadwaj, and J. N. Chengalur. Foregrounds for redshifted 21-cm studies of reionization: Giant Meter Wave Radio Telescope 153-MHz observations. *MNRAS*, 385:2166–2174, April 2008.

[5] Z. S. Ali, A. R. Parsons, H. Zheng, J. C. Pober, A. Liu, J. E. Aguirre, R. F. Bradley, G. Bernardi, C. L. Carilli, C. Cheng, D. R. DeBoer, M. R. Dexter, J. Grobbelaar, J. Horrell, D. C. Jacobs, P. Klima, D. H. E. MacMahon, M. Maree, D. F. Moore, N. Razavi, I. I. Stefan, W. P. Walbrugh, and A. Walker. PAPER-64 Constraints on Reionization: The 21 cm Power Spectrum at z = 8.4. *ApJ*, 809:61, August 2015.

[6] K. M. B. Asad, L. V. E. Koopmans, V. Jelić, A. G. de Bruyn, V. N. Pandey, and B. K. Gehlot. Polarization leakage in epoch of reionization windows - III. Wide-field effects of narrow-field arrays. *MNRAS*, 476(3):3051–3062, May 2018.

[7] K. M. B. Asad, L. V. E. Koopmans, V. Jelić, A. Ghosh, F. B. Abdalla, M. A. Brentjens, A. G. de Bruyn, B. Ciardi, B. K. Gehlot, and I. T. Iliev. Polarization leakage in epoch of reionization windows - II. Primary beam model and direction-dependent calibration. *MNRAS*, 462(4):4482–4494, Nov 2016.

[8] K. M. B. Asad, L. V. E. Koopmans, V. Jelić, V. N. Pand ey, A. Ghosh, F. B. Abdalla, G. Bernardi, M. A. Brentjens, A. G. de Bruyn, and S. Bus. Polarization leakage in epoch of reionization windows - I. Low Frequency Array observations of the 3C196 field. *MNRAS*, 451(4):3709–3727, Aug 2015.

[9] N. Barry, A. P. Beardsley, R. Byrne, B. Hazelton, M. F. Morales, J. C. Pober, and I. Sullivan. The FHD/εppsilon Epoch of Reionization Power Spectrum Pipeline. *arXiv e-prints*, page arXiv:1901.02980, Jan 2019.

[10] N. Barry, B. Hazelton, I. Sullivan, M. F. Morales, and J. C. Pober. Calibration Requirements for Detecting the 21 cm Epoch of Reionization Power Spectrum and Implications for the SKA. *ArXiv e-prints 1603.00607*, March 2016.

[11] A. P. Beardsley, B. J. Hazelton, M. F. Morales, W. Arcus, and others. The EoR sensitivity of the Murchison Widefield Array. *MNRAS*, 429:L5–L9, February 2013.

[12] A. P. Beardsley, B. J. Hazelton, I. S. Sullivan, P. Carroll, N. Barry, M. Rahimi, B. Pindor, C. M. Trott, J. Line, D. C. Jacobs, M. F. Morales, J. C. Pober, G. Bernardi, J. D. Bowman, M. P. Busch, F. Briggs, R. J. Cappallo, B. E. Corey, A. de Oliveira-Costa, J. S. Dillon, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H. S. Kim, E. Kratzenberg, E. Lenc, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, S. J. Tingay, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. First Season MWA EoR Power Spectrum Results at Redshift 7. *ArXiv e-prints*, August 2016.

[13] G. Bernardi, L. J. Greenhill, D. A. Mitchell, S. M. Ord, B. J. Hazelton, B. M. Gaensler, A. de Oliveira-Costa, M. F. Morales, N. Udaya Shankar, and R. Subrahmanyan. A 189 MHz, 2400 deg^2 Polarization Survey with the Murchison Widefield Array 32-element Prototype. *ApJ*, 771(2):105, Jul 2013.

[14] S. Bharadwaj and S. K. Sethi. HI Fluctuations at Large Redshifts: I–Visibility correlation. *Journal of Astrophysics and Astronomy*, 22:293–307, Dec 2001.

[15] Somnath Bharadwaj and Sk. Saiyad Ali. On using visibility correlations to probe the HI distribution from the dark ages to the present epoch - I. Formalism and the expected signal. *MNRAS*, 356(4):1519–1528, Feb 2005.

[16] Somnath Bharadwaj, Srijita Pal, Samir Choudhuri, and Prasun Dutta. A Tapered Gridded Estimator (TGE) for the multifrequency angular power spectrum (MAPS) and the cosmological H I 21-cm power spectrum. *MNRAS*, 483(4):5694–5700, Mar 2019.

[17] J. D. Bowman, I. Cairns, D. L. Kaplan, T. Murphy, D. Oberoi, and others. Science with the Murchison Widefield Array. *PASA*, 30:31, April 2013.

[18] J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and N. Mahesh. An absorption profile centred at 78 megahertz in the sky-averaged spectrum. *Nat*, 555:67–70, March 2018.

[19] Judd D. Bowman, Miguel F. Morales, and Jacqueline N. Hewitt. Foreground Contamination in Interferometric Measurements of the Redshifted 21 cm Power Spectrum. *ApJ*, 695(1):183–199, Apr 2009.

[20] Ruby Byrne, Miguel F. Morales, Bryna Hazelton, Wenyang Li, Nichole Barry, Adam P. Beardsley, Ronniy Joseph, Jonathan Pober, Ian Sullivan, and Cathryn Trott. Fundamental Limitations on the Calibration of Redundant 21 cm Cosmology Instruments and Implications for HERA and the SKA. *ApJ*, 875(1):70, Apr 2019.

[21] P. A. Carroll, J. Line, M. F. Morales, N. Barry, A. P. Beardsley, B. J. Hazelton, D. C. Jacobs, J. C. Pober, I. S. Sullivan, and R. L. Webster. A high reliability survey of discrete Epoch of Reionization foreground sources in the MWA EoR0 field. *MNRAS*, 461(4):4151–4175, Oct 2016.

[22] E. Chapman, S. Zaroubi, and F. Abdalla. Foreground Removal vs. Foreground Avoidance: Contamination of the EoR Window. *ArXiv/1408.4695*, August 2014.

[23] Emma Chapman, Filipe B. Abdalla, J. Bobin, J. L. Starck, Geraint Harker, Vibor Jelić, Panagiotis Labropoulos, Saleem Zaroubi, Michiel A. Brentjens, and A. G. de Bruyn. The scale of the problem: recovering images of reionization with Generalized Morphological Component Analysis. *MNRAS*, 429(1):165–176, Feb 2013.

[24] Emma Chapman, Filipe B. Abdalla, Geraint Harker, Vibor Jelić, Panagiotis Labropoulos, Saleem Zaroubi, Michiel A. Brentjens, A. G. de Bruyn, and L. V. E. Koopmans. Foreground removal using FASTICA: a showcase of LOFAR-EoR. *MNRAS*, 423(3):2518–2532, Jul 2012.

[25] Carina Cheng, Aaron R. Parsons, Matthew Kolopanis, Daniel C. Jacobs, Adrian Liu, Saul A. Kohn, James E. Aguirre, Jonathan C. Pober, Zaki S. Ali, Gianni Bernardi, Richard F. Bradley, Chris L. Carilli, David R. DeBoer, Matthew R. Dexter, Joshua S. Dillon, Pat Klima, David H. E. MacMahon, David F. Moore, Chuneeta D. Nunhokee, William P. Walbrugh, and Andre Walker. Characterizing Signal Loss in the 21 cm Reionization Power Spectrum: A Revised Study of PAPER-64. *ArXiv e-prints*, page arXiv:1810.05175, October 2018.

[26] S. Choudhuri, S. Bharadwaj, A. Ghosh, and S. S. Ali. Visibility-based angular power spectrum estimation in low-frequency radio interferometric observations. *MNRAS*, 445:4351–4365, December 2014.

[27] Samir Choudhuri, Somnath Bharadwaj, Suman Chatterjee, Sk. Saiyad Ali, Nirupam Roy, and Abhik Ghosh. The visibility-based tapered gridded estimator (TGE) for the redshifted 21-cm power spectrum. *MNRAS*, 463(4):4093–4107, Dec 2016.

[28] B. Ciardi, P. Labropoulos, A. Maselli, R. Thomas, S. Zaroubi, L. Graziani, J. S. Bolton, G. Bernardi, M. Brentjens, and A. G. de Bruyn. Prospects for detecting the 21 cm forest from the diffuse intergalactic medium with LOFAR. *MNRAS*, 428(2):1755–1765, Jan 2013.

[29] A. Datta, J. D. Bowman, and C. L. Carilli. Bright Source Subtraction Requirements for Redshifted 21 cm Measurements. *ApJ*, 724:526–538, November 2010.

[30] Kanan K. Datta, T. Roy Choudhury, and Somnath Bharadwaj. The multifrequency angular power spectrum of the epoch of reionization 21-cm signal. *MNRAS*, 378(1):119–128, Jun 2007.

[31] D. R. DeBoer, A. R. Parsons, J. E. Aguirre, P. Alexander, Z. S. Ali, A. P. Beardsley, G. Bernardi, J. D. Bowman, R. F. Bradley, C. L. Carilli, C. Cheng, E. de Lera

Acedo, J. S. Dillon, A. Ewall-Wice, G. Fadana, N. Fagnoni, R. Fritz, S. R. Furlanetto, B. Glendenning, B. Greig, J. Grobbelaar, B. J. Hazelton, J. N. Hewitt, J. Hickish, D. C. Jacobs, A. Julius, M. Kariseb, S. A. Kohn, T. Lekalake, A. Liu, A. Loots, D. MacMahon, L. Malan, C. Malgas, M. Maree, N. Mathison, E. Matsetela, A. Mesinger, M. F. Morales, A. R. Neben, N. Patra, S. Pieterse, J. C. Pober, N. Razavi-Ghods, J. Ringuette, J. Robnett, K. Rosie, R. Sell, C. Smith, A. Syce, M. Tegmark, N. Thyagarajan, P. K. G. Williams, and H. Zheng. Hydrogen Epoch of Reionization Array (HERA). *ArXiv e-prints*, June 2016.

[32] P. Dewdney, W. Turner, R. Braun, J. Santander-Vela, M. Waterson, and G-H Tan. SKA System Baseline Design v2. *SKA Document Series*, 2016.

[33] J. S. Dillon, A. Liu, and M. Tegmark. A fast method for power spectrum and foreground analysis for 21 cm cosmology. , 87(4):043005, February 2013.

[34] J. S. Dillon, A. R. Neben, J. N. Hewitt, M. Tegmark, N. Barry, A. P. Beardsley, J. D. Bowman, F. Briggs, P. Carroll, A. de Oliveira-Costa, A. Ewall-Wice, L. Feng, L. J. Greenhill, B. J. Hazelton, L. Hernquist, N. Hurley-Walker, D. C. Jacobs, H. S. Kim, P. Kittiwisit, E. Lenc, J. Line, A. Loeb, B. McKinley, D. A. Mitchell, M. F. Morales, A. R. Offringa, S. Paul, B. Pindor, J. C. Pober, P. Procopio, J. Riding, S. Sethi, N. U. Shankar, R. Subrahmanyan, I. Sullivan, N. Thyagarajan, S. J. Tingay, C. Trott, R. B. Wayth, R. L. Webster, S. Wyithe, G. Bernardi, R. J. Cappallo, A. A. Deshpande, M. Johnston-Hollitt, D. L. Kaplan, C. J. Lonsdale, S. R. McWhirter, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, K. S. Srivani, A. Williams, and C. L. Williams. Empirical covariance modeling for 21 cm power spectrum estimation: A method demonstration and new limits from early Murchison Widefield Array 128-tile data. , 91(12):123011, June 2015.

[35] D. J. Eisenstein and W. Hu. Power Spectra for Cold Dark Matter and Its Variants. *ApJ*, 511:5–15, January 1999.

[36] S. W. Ellingson, T. E. Clarke, A. Cohen, J. Craig, N. E. Kassim, Y. Pihlstrom, L. J. Rickard, and G. B. Taylor. The Long Wavelength Array. *IEEE Proceedings*, 97:1421–1430, August 2009.

[37] A. Ewall-Wice, Joshua S. Dillon, J. N. Hewitt, A. Loeb, A. Mesinger, A. R. Neben, A. R. Offringa, M. Tegmark, N. Barry, and A. P. Beardsley. First limits on the 21 cm power spectrum during the Epoch of X-ray heating. *MNRAS*, 460(4):4320–4347, Aug 2016.

[38] Aaron Ewall-Wice, Joshua S. Dillon, Adrian Liu, and Jacqueline Hewitt. The impact of modelling errors on interferometer calibration for 21 cm power spectra. *MNRAS*, 470(2):1849–1870, Sep 2017.

[39] S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Phys. Rep.*, 433:181–301, October 2006.

[40] B. K. Gehlot, F. G. Mertens, L. V. E. Koopmans, M. A. Brentjens, S. Zaroubi, B. Ciardi, A. Ghosh, M. Hatef, I. T. Iliev, and V. Jelić. The first power spectrum limit on the 21-cm signal of neutral hydrogen during the Cosmic Dawn at $z = 20 - 25$ from LOFAR. *arXiv e-prints*, page arXiv:1809.06661, Sep 2018.

[41] Abhik Ghosh, Somnath Bharadwaj, Sk. Saiyad Ali, and Jayaram N. Chengalur. GMRT observation towards detecting the post-reionization 21-cm signal. *MNRAS*, 411(4):2426–2438, Mar 2011.

[42] Abhik Ghosh, Somnath Bharadwaj, Sk. Saiyad Ali, and Jayaram N. Chengalur. Improved foreground removal in GMRT 610 MHz observations towards redshifted 21-cm tomography. *MNRAS*, 418(4):2584–2589, Dec 2011.

[43] Abhik Ghosh, Léon V. E. Koopmans, E. Chapman, and V. Jelić. A Bayesian analysis of redshifted 21-cm H I signal and foregrounds: simulations for LOFAR. *MNRAS*, 452(2):1587–1600, Sep 2015.

[44] Abhik Ghosh, Jayanti Prasad, Somnath Bharadwaj, Sk. Saiyad Ali, and Jayaram N. Chengalur. Characterizing foreground for redshifted 21 cm radiation: 150 MHz Giant Metrewave Radio Telescope observations. *MNRAS*, 426(4):3295–3314, Nov 2012.

[45] Bradley Greig, Andrei Mesinger, and Jonathan C. Pober. Constraints on the temperature of the intergalactic medium at z = 8.4 with 21-cm observations. *MNRAS*, 455(4):4295–4300, Feb 2016.

[46] Geraint Harker, Saleem Zaroubi, Gianni Bernardi, Michiel A. Brentjens, A. G. de Bruyn, Benedetta Ciardi, Vibor Jelić, Leon V. E. Koopmans, Panagiotis Labropoulos, and Garrelt Mellema. Power spectrum extraction for redshifted 21-cm Epoch of Reionization experiments: the LOFAR case. *MNRAS*, 405(4):2492–2504, Jul 2010.

[47] Geraint J. A. Harker, Saleem Zaroubi, Rajat M. Thomas, Vibor Jelić, Panagiotis Labropoulos, Garrelt Mellema, Ilian T. Iliev, Gianni Bernardi, Michiel A. Brentjens, and A. G. de Bruyn. Detection and extraction of signals from the epoch of reionization using higher-order one-point statistics. *MNRAS*, 393(4):1449–1458, Mar 2009.

[48] N. Hurley-Walker, J. R. Callingham, P. J. Hancock, T. M. O. Franzen, L. Hindson, A. D. Kapińska, J. Morgan, A. R. Offringa, R. B. Wayth, C. Wu, Q. Zheng, T. Murphy, M. E. Bell, K. S. Dwarakanath, B. For, B. M. Gaensler, M. Johnston-Hollitt, E. Lenc, P. Procopio, L. Staveley-Smith, R. Ekers, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, L. Greenhill, B. J. Hazelton, D. L. Kaplan, C. J. Lonsdale, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, R. L. Webster, A. Williams, and C. L. Williams. GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey - I. A low-frequency extragalactic catalogue. *MNRAS*, 464:1146–1167, January 2017.

[49] Natasha Hurley-Walker, John Morgan, Randall B. Wayth, Paul J. Hancock, Martin E. Bell, Gianni Bernardi, Ramesh Bhat, Frank Briggs, Avinash A. Deshpande,

and Aaron Ewall-Wice. The Murchison Widefield Array Commissioning Survey: A Low-Frequency Catalogue of 14 110 Compact Radio Sources over 6 100 Square Degrees. *PASA*, 31:e045, Nov 2014.

[50] Anne Hutter, Pratika Dayal, Sangeeta Malhotra, James Rhoads, Tirthankar Roy Choudhury, Benedetta Ciardi, Christopher J. Conselice, Asantha Cooray, Jean-Gabriel Cuby, and Kanan K. Datta. Astro2020 Science White Paper: A proposal to exploit galaxy-21cm synergies to shed light on the Epoch of Reionization. *arXiv e-prints*, page arXiv:1903.03628, Mar 2019.

[51] Anne Hutter, Pratika Dayal, Volker Müller, and Cathryn M. Trott. Exploring 21cm-Lyman Alpha Emitter Synergies for SKA. *ApJ*, 836(2):176, Feb 2017.

[52] D. C. Jacobs, B. J. Hazelton, C. M. Trott, J. S. Dillon, B. Pindor, I. S. Sullivan, J. C. Pober, N. Barry, A. P. Beardsley, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H. Kim, E. Kratzenberg, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, S. J. Tingay, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. The Murchison Widefield Array 21 cm Power Spectrum Analysis Methodology. *ApJ*, 825:114, July 2016.

[53] Daniel C. Jacobs, James E. Aguirre, Aaron R. Parsons, Jonathan C. Pober, Richard F. Bradley, Chris L. Carilli, Nicole E. Gugliucci, Jason R. Manley, Carel van der Merwe, and David F. Moore. New 145 MHz Source Measurements by PAPER in the Southern Sky. *ApJ*, 734(2):L34, Jun 2011.

[54] Daniel C. Jacobs, Jacob Burba, Judd D. Bowman, Abraham R. Neben, Benjamin Stinnett, Lauren Turner, Kali Johnson, Michael Busch, Jay Allison, and Marc Leatham. First Demonstration of ECHO: an External Calibrator for Hydrogen Observatories. *PASP*, 129(973):035002, Mar 2017.

[55] Daniel C. Jacobs, Jonathan C. Pober, Aaron R. Parsons, James E. Aguirre, Zaki S. Ali, Judd Bowman, Richard F. Bradley, Chris L. Carilli, David R. DeBoer, and Matthew R. Dexter. Multiredshift Limits on the 21 cm Power Spectrum from PAPER. *ApJ*, 801(1):51, Mar 2015.

[56] V. Jelić, A. G. de Bruyn, M. Mevius, F. B. Abdalla, K. M. B. Asad, G. Bernardi, M. A. Brentjens, S. Bus, E. Chapman, and B. Ciardi. Initial LOFAR observations of epoch of reionization windows. II. Diffuse polarized emission in the ELAIS-N1 field. *A&A*, 568:A101, Aug 2014.

[57] V. Jelić, S. Zaroubi, P. Labropoulos, R. M. Thomas, G. Bernardi, M. A. Brentjens, A. G. de Bruyn, B. Ciardi, G. Harker, and L. V. E. Koopmans. Foreground simulations

for the LOFAR-epoch of reionization experiment. *MNRAS*, 389(3):1319–1335, Sep 2008.

[58] Vibor Jelić, Saleem Zaroubi, Panagiotis Labropoulos, Gianni Bernardi, A. G. de Bruyn, and Léon V. E. Koopmans. Realistic simulations of the Galactic polarized foreground: consequences for 21-cm reionization detection experiments. *MNRAS*, 409(4):1647–1659, Dec 2010.

[59] C. H. Jordan, S. Murray, C. M. Trott, R. B. Wayth, D. A. Mitchell, M. Rahimi, B. Pindor, P. Procopio, and J. Morgan. Characterization of the ionosphere above the Murchison Radio Observatory using the Murchison Widefield Array. *MNRAS*, 471:3974–3987, November 2017.

[60] Ronniy C. Joseph, Cathryn M. Trott, and Randall B. Wayth. The Bias and Uncertainty of Redundant and Sky-based Calibration Under Realistic Sky and Telescope Conditions. *AJ*, 156(6):285, Dec 2018.

[61] S. Kazemi, S. Yatawatta, S. Zaroubi, P. Lampropoulos, A. G. de Bruyn, L. V. E. Koopmans, and J. Noordam. Radio interferometric calibration using the SAGE algorithm. *MNRAS*, 414(2):1656–1666, Jun 2011.

[62] Joshua Kerrigan, Paul La Plante, Saul Kohn, Jonathan C. Pober, James Aguirre, Zara Abdurashidova, Paul Alexander, Zaki S. Ali, Yanga Balfour, and Adam P. Beardsley. Optimizing Sparse RFI Prediction using Deep Learning. *arXiv e-prints*, page arXiv:1902.08244, Feb 2019.

[63] Joshua R. Kerrigan, Jonathan C. Pober, Zaki S. Ali, Carina Cheng, Adam P. Beardsley, Aaron R. Parsons, James E. Aguirre, Nichole Barry, Richard F. Bradley, and Gianni Bernardi. Improved 21 cm Epoch of Reionization Power Spectrum Measurements with a Hybrid Foreground Subtraction and Avoidance Technique. *ApJ*, 864(2):131, Sep 2018.

[64] S. A. Kohn, J. E. Aguirre, C. D. Nunhokee, G. Bernardi, J. C. Pober, Z. S. Ali, R. F. Bradley, C. L. Carilli, D. R. DeBoer, and N. E. Gugliucci. Constraining Polarized Foregrounds for EoR Experiments I: 2D Power Spectra from the PAPER-32 Imaging Array. *ApJ*, 823(2):88, Jun 2016.

[65] L. Koopmans, J. Pritchard, G. Mellema, J. Aguirre, K. Ahn, R. Barkana, I. van Bemmel, G. Bernardi, A. Bonaldi, F. Briggs, and others. The Cosmic Dawn and Epoch of Reionisation with SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 1, 2015.

[66] Kenji Kubota, Shintaro Yoshiura, Keitaro Takahashi, Kenji Hasegawa, Hidenobu Yajima, Masami Ouchi, B. Pindor, and R. L. Webster. Detectability of the 21-cm signal during the epoch of reionization with 21-cm Lyman α emitter cross-correlation - I. *MNRAS*, 479(2):2754–2766, Sep 2018.

[67] Adam Lanman, Bryna Hazelton, Daniel Jacobs, Matthew Kolopanis, Jonathan Pober, James Aguirre, and Nithyanandan Thyagarajan. pyuvsim: A comprehensive simulation package for radio interferometers in python. *The Journal of Open Source Software*, 4(37):1234, May 2019.

[68] E. Lenc, C. S. Anderson, N. Barry, J. D. Bowman, I. H. Cairns, J. S. Farnes, B. M. Gaensler, G. Heald, M. Johnston-Hollitt, and D. L. Kaplan. The Challenges of Low-Frequency Radio Polarimetry: Lessons from the Murchison Widefield Array. *PASA*, 34:e040, Sep 2017.

[69] E. Lenc, B. M. Gaensler, X. H. Sun, E. M. Sadler, A. G. Willis, N. Barry, A. P. Beardsley, M. E. Bell, G. Bernardi, and J. D. Bowman. Low-frequency Observations of Linearly Polarized Structures in the Interstellar Medium near the South Galactic Pole. *ApJ*, 830(1):38, Oct 2016.

[70] W. Li, J. C. Pober, B. J. Hazelton, N. Barry, M. F. Morales, I. Sullivan, A. R. Parsons, Z. S. Ali, J. S. Dillon, and A. P. Beardsley. Comparing Redundant and Sky-model-based Interferometric Calibration: A First Look with Phase II of the MWA. *ApJ*, 863(2):170, Aug 2018.

[71] J. L. B. Line, B. McKinley, J. Rasti, M. Bhardwaj, R. B. Wayth, R. L. Webster, D. Ung, D. Emrich, L. Horsley, and A. Beardsley. In situ measurement of MWA primary beam variation using ORBCOMM. *PASA*, 35:45, Dec 2018.

[72] J. L. B. Line, R. L. Webster, B. Pindor, D. A. Mitchell, and C. M. Trott. PUMA: The Positional Update and Matching Algorithm. *PASA*, 34:e003, Jan 2017.

[73] A. Liu and M. Tegmark. A method for 21 cm power spectrum estimation in the presence of foregrounds. *PhRvD*, 83(10):103006–+, May 2011.

[74] Adrian Liu, Max Tegmark, Scott Morrison, Andrew Lutomirski, and Matias Zaldarriaga. Precision calibration of radio interferometers using redundant baselines. *MNRAS*, 408(2):1029–1050, Oct 2010.

[75] Zachary E. Martinot, James E. Aguirre, Saul A. Kohn, and Immanuel Q. Washington. Ionospheric Attenuation of Polarized Foregrounds in 21 cm Epoch of Reionization Measurements: A Demonstration for the HERA Experiment. *ApJ*, 869(1):79, Dec 2018.

[76] M. McQuinn, O. Zahn, M. Zaldarriaga, L. Hernquist, and S. R. Furlanetto. Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization. *ApJ*, 653:815–834, December 2006.

[77] F. G. Mertens, A. Ghosh, and L. V. E. Koopmans. Statistical 21-cm signal separation via Gaussian Process Regression analysis. *MNRAS*, 478(3):3640–3652, Aug 2018.

[78] A. Mesinger, S. Furlanetto, and R. Cen. 21CMFAST: a fast, seminumerical simulation of the high-redshift 21-cm signal. *MNRAS*, 411:955–972, February 2011.

[79] D. A. Mitchell, L. J. Greenhill, R. B. Wayth, R. J. Sault, C. J. Lonsdale, R. J. Cappallo, M. F. Morales, and S. M. Ord. Real-Time Calibration of the Murchison Widefield Array. *IEEE Journal of Selected Topics in Signal Processing, Vol. 2, Issue 5, p.707-717*, 2:707–717, November 2008.

[80] D. F. Moore, J. E. Aguirre, A. R. Parsons, D. C. Jacobs, and J. C. Pober. The Effects of Polarized Foregrounds on 21 cm Epoch of Reionization Power Spectrum Measurements. *ApJ*, 769:154, June 2013.

[81] M. F. Morales and J. Hewitt. Toward Epoch of Reionization Measurements with Wide-Field Radio Observations. *ApJ*, 615:7–18, November 2004.

[82] Miguel F. Morales, Adam Beardsley, Jonathan Pober, Nichole Barry, Bryna Hazelton, Daniel Jacobs, and Ian Sullivan. Understanding the diversity of 21 cm cosmology analyses. *MNRAS*, 483(2):2207–2216, Feb 2019.

[83] A. Mouri Sardarabadi and L. V. E. Koopmans. Quantifying suppression of the cosmological 21-cm signal due to direction-dependent gain calibration in radio interferometers. *MNRAS*, 483(4):5480–5490, Mar 2019.

[84] S. G. Murray, C. M. Trott, and C. H. Jordan. An Improved Statistical Point-source Foreground Model for the Epoch of Reionization. *ApJ*, 845:7, August 2017.

[85] A. R. Neben, R. F. Bradley, J. N. Hewitt, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, R. Goeke, and L. J. Greenhill. Measuring phased-array antenna beampatterns with high dynamic range for the Murchison Widefield Array using 137 MHz ORBCOMM satellites. *Radio Science*, 50(7):614–629, Jul 2015.

[86] Abraham R. Neben, Richard F. Bradley, Jacqueline N. Hewitt, David R. DeBoer, Aaron R. Parsons, James E. Aguirre, Zaki S. Ali, Carina Cheng, Aaron Ewall-Wice, and Nipanjana Patra. The Hydrogen Epoch of Reionization Array Dish. I. Beam Pattern Measurements and Science Implications. *ApJ*, 826(2):199, Aug 2016.

[87] C. D. Nunhokee, G. Bernardi, S. A. Kohn, J. E. Aguirre, N. Thyagarajan, J. S. Dillon, G. Foster, T. L. Grobler, J. Z. E. Martinot, and A. R. Parsons. Constraining Polarized Foregrounds for EoR Experiments. II. Polarization Leakage Simulations in the Avoidance Scheme. *ApJ*, 848(1):47, Oct 2017.

[88] A. R. Offringa, F. Mertens, and L. V. E. Koopmans. The impact of interference excision on 21-cm epoch of reionization power spectrum analyses. *MNRAS*, 484(2):2866–2875, Apr 2019.

[89] A. R. Offringa, F. Mertens, and L. V. E. Koopmans. The impact of interference excision on 21-cm epoch of reionization power spectrum analyses. *MNRAS*, 484(2):2866–2875, Apr 2019.

[90] A. R. Offringa, J. J. van de Gronde, and J. B. T. M. Roerdink. A morphological algorithm for improved radio-frequency interference detection. *A&A*, 539, March 2012.

[91] Gregory Paciga, Joshua G. Albert, Kevin Bandura, Tzu-Ching Chang, Yashwant Gupta, Christopher Hirata, Julia Odegova, Ue-Li Pen, Jeffrey B. Peterson, Jayanta Roy, J. Richard Shaw, Kris Sigurdson, and Tabitha Voytek. A simulation-calibrated limit on the H I power spectrum from the GMRT Epoch of Reionization experiment. *MNRAS*, 433(1):639–647, Jul 2013.

[92] Gregory Paciga, Tzu-Ching Chang, Yashwant Gupta, Rajaram Nityanada, Julia Odegova, Ue-Li Pen, Jeffrey B. Peterson, Jayanta Roy, and Kris Sigurdson. The GMRT Epoch of Reionization experiment: a new upper limit on the neutral hydrogen power spectrum at z≈ 8.6. *MNRAS*, 413(2):1174–1183, May 2011.

[93] A. Parsons, J. Pober, M. McQuinn, D. Jacobs, and J. Aguirre. A Sensitivity and Array-configuration Study for Measuring the Power Spectrum of 21 cm Emission from Reionization. *ApJ*, 753:81, July 2012.

[94] A. R. Parsons, D. C. Backer, G. S. Foster, M. C. H. Wright, R. F. Bradley, N. E. Gugliucci, C. R. Parashare, E. E. Benoit, J. E. Aguirre, D. C. Jacobs, C. L. Carilli, D. Herne, M. J. Lynch, J. R. Manley, and D. J. Werthimer. The Precision Array for Probing the Epoch of Re-ionization: Eight Station Results. *AJ*, 139:1468–1480, April 2010.

[95] A. R. Parsons, A. Liu, J. E. Aguirre, Z. S. Ali, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, D. C. Jacobs, P. Klima, D. H. E. MacMahon, J. R. Manley, D. F. Moore, J. C. Pober, I. I. Stefan, and W. P. Walbrugh. New Limits on 21 cm Epoch of Reionization from PAPER-32 Consistent with an X-Ray Heated Intergalactic Medium at z = 7.7. *ApJ*, 788:106, June 2014.

[96] Aaron Parsons, Donald Backer, Andrew Siemion, Henry Chen, Dan Werthimer, Pierre Droz, Terry Filiba, Jason Manley, Peter McMahon, and Arash Parsa. A Scalable Correlator Architecture Based on Modular FPGA Hardware, Reuseable Gateware, and Data Packetization. *PASP*, 120(873):1207, Nov 2008.

[97] Aaron R. Parsons, Adrian Liu, Zaki S. Ali, and Carina Cheng. Optimized Beam Sculpting with Generalized Fringe-rate Filters. *ApJ*, 820(1):51, Mar 2016.

[98] Aaron R. Parsons, Jonathan C. Pober, James E. Aguirre, Christopher L. Carilli, Daniel C. Jacobs, and David F. Moore. A Per-baseline, Delay-spectrum Technique for Accessing the 21 cm Cosmic Reionization Signature. *ApJ*, 756(2):165, Sep 2012.

[99] A. H. Patil, S. Yatawatta, L. V. E. Koopmans, A. G. de Bruyn, M. A. Brentjens, S. Zaroubi, K. M. B. Asad, M. Hatef, V. Jelić, and M. Mevius. Upper Limits on the 21 cm Epoch of Reionization Power Spectrum from One Night with LOFAR. *ApJ*, 838(1):65, Mar 2017.

[100] A. H. Patil, S. Yatawatta, S. Zaroubi, L. V. E. Koopmans, A. G. de Bruyn, V. Jelić, B. Ciardi, I. T. Iliev, M. Mevius, V. N. Pandey, and B. K. Gehlot. Systematic biases in low frequency radio interferometric data due to calibration: the LOFAR EoR case. *MNRAS*, September 2016.

[101] A. H. Patil, S. Zaroubi, E. Chapman, V. Jelić, G. Harker, F. B. Abdalla, K. M. B. Asad, G. Bernardi, M. A. Brentjens, A. G. de Bruyn, S. Bus, B. Ciardi, S. Daiboo, E. R. Fernandez, A. Ghosh, H. Jensen, S. Kazemi, L. V. E. Koopmans, P. Labropoulos, M. Mevius, O. Martinez, G. Mellema, A. R. Offringa, V. N. Pandey, J. Schaye, R. M. Thomas, H. K. Vedantham, V. Veligatla, S. J. Wijnholds, and S. Yatawatta. Constraining the epoch of reionization with the variance statistic: simulations of the LOFAR case. *MNRAS*, 443:1113–1124, September 2014.

[102] Ajinkya H. Patil, Saleem Zaroubi, Emma Chapman, Vibor Jelić, Geraint Harker, Filipe B. Abdalla, Khan M. B. Asad, Gianni Bernardi, Michiel A. Brentjens, and A. G. de Bruyn. Constraining the epoch of reionization with the variance statistic: simulations of the LOFAR case. *MNRAS*, 443(2):1113–1124, Sep 2014.

[103] J. C. Pober, A. Liu, J. S. Dillon, J. E. Aguirre, J. D. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, J. N. Hewitt, D. C. Jacobs, M. McQuinn, M. F. Morales, A. R. Parsons, M. Tegmark, and D. J. Werthimer. What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization. *ApJ*, 782:66, February 2014.

[104] Jonathan C. Pober, Zaki S. Ali, Aaron R. Parsons, Matthew McQuinn, James E. Aguirre, Gianni Bernardi, Richard F. Bradley, Chris L. Carilli, Carina Cheng, and David R. DeBoer. PAPER-64 Constraints On Reionization. II. The Temperature of the z =8.4 Intergalactic Medium. *ApJ*, 809(1):62, Aug 2015.

[105] Jonathan C. Pober, Aaron R. Parsons, Daniel C. Jacobs, James E. Aguirre, Richard F. Bradley, Chris L. Carilli, Nicole E. Gugliucci, David F. Moore, and Chaitali R. Parashare. A Technique for Primary Beam Calibration of Drift-scanning, Wide-field Antenna Elements. *AJ*, 143(2):53, Feb 2012.

[106] P. Procopio, B. Pindor, D.A.M. Mitchell, and builders. Calibration of MWA data using the MWA Real-Time System. *ApJ*, 2015, in prep.

[107] Sandeep Rana and Jasjeet S. Bagla. Angular clustering of point sources at 150 MHz in the TGSS survey. *MNRAS*, 485(4):5891–5896, Jun 2019.

[108] SK. Saiyad Ali, Somnath Bharadwaj, and Sanjay K. Pandey. Probing the bispectrum at high redshifts using 21-cm HI observations. *MNRAS*, 366(1):213–218, Feb 2006.

[109] Irina I. Stefan, Chris L. Carilli, David A. Green, Zaki Ali, James E. Aguirre, Richard F. Bradley, David DeBoer, Matthew Dexter, Nicole E. Gugliucci, and D. E. Harris. Imaging on PAPER: Centaurus A at 148 MHz. *MNRAS*, 432(2):1285–1293, Jun 2013.

[110] I. S. Sullivan, M. F. Morales, B. J. Hazelton, W. Arcus, D. Barnes, G. Bernardi, F. H. Briggs, J. D. Bowman, J. D. Bunton, and R. J. Cappallo. Fast Holographic Deconvolution: A New Technique for Precision Radio Interferometry. *ApJ*, 759(1):17, Nov 2012.

[111] G. Swarup, S. Ananthakrishnan, V. K. Kapahi, A. P. Rao, C. R. Subrahmanya, and V. K. Kulkarni. The Giant Metre-Wave Radio Telescope. *Current Science*, 60:95, Jan 1991.

[112] M. Tegmark. How to measure CMB power spectra without losing information. *PhRvD*, 55:5895–5907, May 1997.

[113] A. R. Thompson, J. M. Moran, and G. W. Swenson. *Interferometry and synthesis in radio astronomy*. 1986.

[114] N. Thyagarajan, D. C. Jacobs, J. D. Bowman, and others. Foregrounds in Wide-field Redshifted 21 cm Power Spectra. *ApJ*, 804:14, May 2015.

[115] S. J. Tingay, R. Goeke, J. D. Bowman, D. Emrich, and others. The Murchison Wide-field Array: The Square Kilometre Array Precursor at Low Radio Frequencies. *PASA*, 30:7, January 2013.

[116] C. M. Trott, B. Pindor, P. Procopio, R. B. Wayth, D. A. Mitchell, B. McKinley, S. J. Tingay, N. Barry, A. P. Beardsley, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, A. de Oliveira-Costa, J. S. Dillon, A. Ewall-Wice, L. Feng, L. J. Greenhill, B. J. Hazelton, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. C. Jacobs, D. L. Kaplan, H. S. Kim, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. F. Morales, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, J. C. Pober, T. Prabu, J. Riding, N. Udaya Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, R. L. Webster, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. CHIPS: The Cosmological H i Power Spectrum Estimator. *ApJ*, 818:139, February 2016.

[117] C. M. Trott and R. B. Wayth. Spectral Calibration Requirements of Radio Interferometers for Epoch of Reionisation Science with the SKA. *PASA*, 33:e019, May 2016.

[118] C. M. Trott, R. B. Wayth, and S. J. Tingay. The Impact of Point-source Subtraction Residuals on 21 cm Epoch of Reionization Estimation. *ApJ*, 757:101, September 2012.

[119] Cathryn M. Trott, Eloy de Lera Acedo, Randall B. Wayth, Nicolas Fagnoni, Adrian T. Sutinjo, Brett Wakley, and Chris Ivan B. Punzalan. Spectral performance of Square Kilometre Array Antennas - II. Calibration performance. *MNRAS*, 470(1):455–465, Sep 2017.

[120] Cathryn M. Trott, Shih Ching Fu, S. G. Murray, C. H. Jordan, J. L. B. Line, N. Barry, R. Byrne, B. J. Hazelton, K. Hasegawa, and R. Joseph. Robust statistics towards detection of the 21 cm signal from the Epoch of Reionization. *MNRAS*, 486(4):5766–5784, Jul 2019.

[121] Cathryn M. Trott, C. H. Jordan, S. G. Murray, B. Pindor, D. A. Mitchell, R. B. Wayth, J. Line, B. McKinley, A. Beardsley, and J. Bowman. Assessment of Ionospheric Activity Tolerances for Epoch of Reionization Science with the Murchison Widefield Array. *ApJ*, 867(1):15, Nov 2018.

[122] Cathryn M. Trott, Catherine A. Watkinson, Christopher H. Jordan, Shintaro Yoshiura, Suman Majumdar, N. Barry, R. Byrne, B. J. Hazelton, K. Hasegawa, and R. Joseph. Gridded and direct Epoch of Reionisation bispectrum estimates using the Murchison Widefield Array. *arXiv e-prints*, page arXiv:1905.07161, May 2019.

[123] Cathryn M. Trott and Randall B. Wayth. Building Models for Extended Radio Sources: Implications for Epoch of Reionisation Science. *PASA*, 34:e061, Nov 2017.

[124] M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, J. P. McKean, J. W. T. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, M. Brentjens, A. Nelles, R. Beck, H. Falcke, R. Fender, J. Hörandel, L. V. E. Koopmans, G. Mann, G. Miley, H. Röttgering, B. W. Stappers, and others. LOFAR: The LOw-Frequency ARray. *A&A*, 556:A2, August 2013.

[125] R. B. Wayth, E. Lenc, M. E. Bell, J. R. Callingham, K. S. Dwarakanath, T. M. O. Franzen, B.-Q. For, B. Gaensler, P. Hancock, L. Hindson, N. Hurley-Walker, C. A. Jackson, M. Johnston-Hollitt, A. D. Kapińska, B. McKinley, J. Morgan, A. R. Offringa, P. Procopio, L. Staveley-Smith, C. Wu, Q. Zheng, C. M. Trott, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, B. E. Corey, A. A. Deshpande, D. Emrich, R. Goeke, L. J. Greenhill, B. J. Hazelton, D. L. Kaplan, J. C. Kasper, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, A. E. E. Rogers, A. Roshi, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and C. L. Williams. GLEAM: The GaLactic and Extragalactic All-Sky MWA Survey. *PASA*, 32:e025, June 2015.

[126] R. B. Wayth, S. J. Tingay, C. M. Trott, D. Emrich, M. Johnston-Hollitt, B. McKinley, B. M. Gaensler, A. P. Beardsley, T. Booler, B. Crosse, T. M. O. Franzen, L. Horsley, D. L. Kaplan, D. Kenney, M. F. Morales, D. Pallot, G. Sleap, K. Steele, M. Walker, A. Williams, C. Wu, I. H. Cairns, M. D. Filipovic, S. Johnston, T. Murphy, P. Quinn, L. Staveley-Smith, R. Webster, and J. S. B. Wyithe. The Phase II Murchison Widefield Array: Design Overview. *ArXiv e-prints*, September 2018.

[127] Michael J. Wilensky, Miguel F. Morales, Bryna J. Hazelton, Nichole Barry, Ruby Byrne, and Sumit Roy. Absolving the SSINS of Precision Interferometric Radio Data: A New Technique for Mitigating Ultra-faint Radio Frequency Interference. *arXiv e-prints*, page arXiv:1906.01093, Jun 2019.

[128] Christopher L. Williams, Jacqueline N. Hewitt, Alan M. Levine, Angelica de Oliveira-Costa, Judd D. Bowman, Frank H. Briggs, B. M. Gaensler, Lars L. Hernquist, Daniel A. Mitchell, and Miguel F. Morales. Low-frequency Imaging of Fields at High Galactic Latitude with the Murchison Widefield Array 32 Element Prototype. *ApJ*, 755(1):47, Aug 2012.

[129] J. S. B. Wyithe and M. F. Morales. Biased reionization and non-Gaussianity in redshifted 21-cm intensity maps of the reionization epoch. *MNRAS*, 379:1647–1657, August 2007.

[130] S. Yatawatta, A. G. de Bruyn, M. A. Brentjens, P. Labropoulos, V. N. Pandey, S. Kazemi, S. Zaroubi, L. V. E. Koopmans, A. R. Offringa, and V. Jelić. Initial deep LOFAR observations of epoch of reionization windows. I. The north celestial pole. *A&A*, 550:A136, Feb 2013.

[131] S. Yoshiura, K. Ichiki, B. Pindor, K. Takahashi, H. Tashiro, and C. M. Trott. Study of systematics effects on the cross power spectrum of 21 cm line and cosmic microwave background using Murchison Widefield Array data. *MNRAS*, 483(2):2697–2711, Feb 2019.

[132] S. Yoshiura, J. L. B. Line, K. Kubota, K. Hasegawa, and K. Takahashi. Detectability of 21 cm-signal during the Epoch of Reionization with 21 cm-Lyman-α emitter cross-correlation - II. Foreground contamination. *MNRAS*, 479(2):2767–2776, Sep 2018.

[133] Shintaro Yoshiura, Hayato Shimabukuro, Keitaro Takahashi, Rieko Momose, Hiroyuki Nakanishi, and Hiroshi Imai. Sensitivity for 21 cm bispectrum from Epoch of Reionization. *MNRAS*, 451(1):266–274, Jul 2015.

[134] S. Zaroubi, A. G. de Bruyn, G. Harker, R. M. Thomas, P. Labropolous, V. Jelić, L. V. E. Koopmans, M. A. Brentjens, G. Bernardi, and B. Ciardi. Imaging neutral hydrogen on large scales during the Epoch of Reionization with LOFAR. *MNRAS*, 425(4):2964–2973, Oct 2012.

[135] H. Zheng, M. Tegmark, V. Buza, J. S. Dillon, H. Gharibyan, J. Hickish, E. Kunz, A. Liu, J. Losh, and A. Lutomirski. MITEoR: a scalable interferometer for precision 21 cm cosmology. *MNRAS*, 445(2):1084–1103, Dec 2014.

[136] Qian Zheng, Xiang-Ping Wu, Jun-Hua Gu, Jingying Wang, and Haiguang Xu. A Method to Extract the Angular Power Spectrum of the Epoch of Reionization from Low-frequency Radio Interferometers. *ApJ*, 758(1):L24, Oct 2012.

Chapter 9

Chapter title

Author Name

Abstract

This chapter discusses some important things

9.1   A Section

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis eu egestas erat. Maecenas tincidunt lacinia tincidunt. Mauris id lectus nec neque feugiat condimentum vitae at diam. In vel orci nunc, non commodo mauris. Vivamus ipsum enim, vulputate quis pharetra non, molestie quis felis. Vivamus porttitor placerat turpis at accumsan. Nunc tortor velit, faucibus a rhoncus nec, blandit non elit. Nam consectetur lectus eu nisi blandit dapibus rhoncus dui tempus. Mauris fermentum dolor vel ipsum vulputate sit amet ultricies tortor lacinia. Donec ut nibh erat. Morbi nec mi ante. Integer nec vestibulum diam. Donec tincidunt pellentesque quam, ut interdum mauris venenatis condimentum. Nam condimentum, augue in aliquet gravida, neque dui elementum eros, id semper eros purus sed felis. Curabitur in justo sit amet sapien ultrices hendrerit at quis nibh. Quisque iaculis pulvinar tincidunt.

$$
\begin{aligned}
C(12) &= \left[\vec{\pi} \cdot \vec{\phi}\,(x+r) \right] \\
&\approx 1 - \text{const}\frac{r^2}{L^2} \int_r^L \frac{x\mathrm{d}x}{x^2} + \cdots \\
&\approx 1 - \text{const}\frac{r^2}{L^2} \ln \frac{x\mathrm{d}x}{x^2} + \cdots .
\end{aligned}
\tag{9.1}
$$

Aenean tellus risus, porta sit amet porta vitae, tincidunt ut felis. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Phasellus pulvinar placerat velit auctor egestas. Vivamus euismod fringilla tincidunt. Sed ut magna felis, id sollicitudin nunc. Quisque a dui eu erat consectetur egestas a quis justo. Aenean euismod congue diam, vel posuere urna fermentum sit amet. Lorem ipsum dolor sit amet, consectetur adipiscing

Figure 9.1: This is figure 1 in chapter 1.

Table 9.1: Greek Letters.

α	β	γ	δ	ε	ε	ζ	η
θ	ϑ	γ	κ	λ	μ	ν	ξ
o	π	ϖ	ρ	ρ	σ	ς	
τ	υ	ϕ	φ	χ	ψ	ω	
Γ	Δ	Θ	Λ	Ξ	Π	Σ	Υ
Φ	Ψ	Ω					

elit. Mauris faucibus lacus eget est mollis auctor. Donec at nibh ligula, et posuere massa. Phasellus quis leo diam [1]. Donec aliquam blandit risus, eu venenatis ante euismod eu. Curabitur cursus justo id arcu condimentum feugiat. Integer sapien urna, vulputate et adipiscing nec, convallis et justo. Suspendisse in ipsum at felis ornare interdum [2],

Cras adipiscing   sagittis nunc vel luctus. Suspendisse volutpat augue quis erat semper consequat dignissim tellus euismod. Morbi hendrerit, tellus id aliquam iaculis, nibh leo tincidunt eros, vitae varius ligula felis in mi.
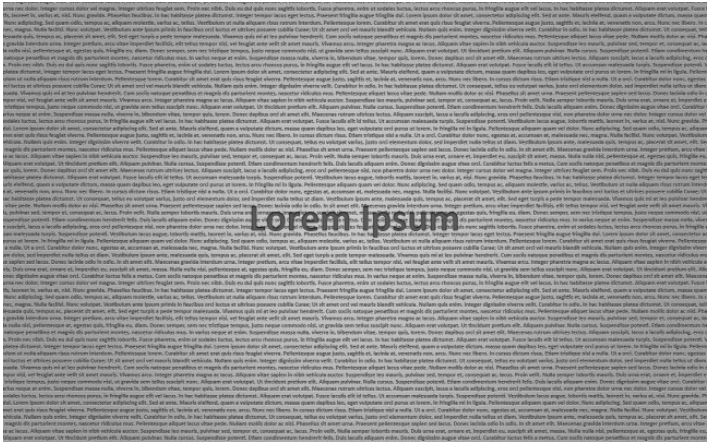
Figure 9.2: This is figure 2 in chapter 1.

Bibliography

[1] KI Diamantaras and SY Kung. *Principal component neural networks: theory and applications*. John Wiley & Sons, Inc. New York, NY, USA, 1996.

[2] D. Tulone and S. Madden. PAQ: Time Series Forecasting for Approximate Query Answering in Sensor Networks. In *Proceedings of the 3rd European Workshop on Wireless Sensor Networks*, pages 21–37. Springer, 2006.