

Generating extremely large-volume reionization simulations

Bradley Greig¹,^{1,2}★ J. Stuart B. Wyithe,^{1,2} Steven G. Murray³,³ Simon J. Mutch^{1,2}
and Cathryn M. Trott^{2,4}

¹*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

²*ARC Centre of Excellence for All-Sky Astrophysics in 3 Dimensions (ASTRO 3D)*

³*School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA*

⁴*International Centre for Radio Astronomy Research (ICRAR), Curtin University, Bentley, WA 6102, Australia*

Accepted 2022 September 1. Received 2022 August 24; in original form 2022 May 20

ABSTRACT

Preparing for a first detection of the 21-cm signal during reionization by large-scale interferometer experiments requires rigorous testing of the data analysis and reduction pipelines. Validating that these do not erroneously add/remove features mimicking the signal (e.g. from side lobes or large-scale power leakage) requires simulations extending beyond the primary field of view. However, the Murchison Wide Field Array (MWA) with a field of view of $\sim 25^2$ deg² would require simulations spanning several Gpcs, which are currently infeasible. To address this, we developed a simplified version of the seminumerical reionization simulation code 21CMFAST, sacrificing some physical accuracy (linear structure formation) in favour of extremely large volumes. We then constructed a 7.5 Gpc comoving volume specifically tailored to the binned spectral resolution of the MWA (~ 1.17 cMpc), required for validating the pipeline used in the 2020 MWA 21-cm power spectrum (PS) upper limits. With this large-volume simulation, we then explored: (i) whether smaller volume simulations are biased by missing large-scale modes, (ii) non-Gaussianity in the cosmic variance uncertainty, (iii) biases in the recovered 21-cm PS following foreground wedge avoidance, and (iv) the impact of tiling smaller simulations to achieve large volumes. We found (i) no biases from missing large-scale power, (ii) significant contribution from non-Gaussianity, as expected, (iii) a 10–20 per cent overestimate of the 21-cm PS following wedge mode excision, and (iv) tiling smaller simulations underestimates the large-scale power and cosmic variance.

Key words: galaxies: high-redshift – cosmology: theory – intergalactic medium – dark ages, reionization, first stars – diffuse radiation – early Universe.

1 INTRODUCTION

Directly observing the first stars and primordial galaxies is rendered nigh on impossible by the inescapable neutral hydrogen fog that enshrouds the early Universe. This fog is gradually lifted in the intergalactic medium (IGM) as the neutral hydrogen is ionized by the cumulative output of ultraviolet (UV) photons from stars and galaxies. This phase transition is referred to as the epoch of reionization (EoR).

Our most promising method to observe the EoR is through detecting the 21-cm hyperfine spin-flip transition of neutral hydrogen. This faint signal, measured relative to the cosmic microwave background (CMB), is in either emission or absorption depending on both the ionization and thermal state of the IGM (see e.g. Gnedin & Ostriker 1997; Madau, Meiksin & Rees 1997; Shaver et al. 1999; Tozzi et al. 2000; Gnedin & Shaver 2004; Furlanetto, Oh & Briggs 2006; Morales & Wyithe 2010; Pritchard & Loeb 2012). Crucially, by measuring the 21-cm signal we gain a 2D picture of the spatial distribution of the neutral hydrogen in the IGM. Being a line transition, repeating this over a broad frequency range builds up

a full 3D time-dependent movie of the IGM in the early Universe. From this, we can infer the properties of the galaxies responsible for driving the EoR.

To detect this 3D signal we require large-scale interferometer experiments, which are specifically designed to be sensitive to the spatial fluctuations. The first generation of these experiments include the Low-Frequency Array (LOFAR; van Haarlem et al. 2013), the Murchison Wide Field Array (MWA; Tingay et al. 2013; Wayth et al. 2018), the Precision Array for Probing the Epoch of Reionisation (Parsons et al. 2010), the Owens Valley Radio Observatory Long Wavelength Array (Eastwood et al. 2019), and the upgraded Giant Metrewave Radio Telescope (Gupta et al. 2017). Next-generation experiments, offering larger collecting areas and lower noise, are currently underway or are planned such as the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al. 2017), New extension in Nançay Upgrading loFAR (Zarka et al. 2012), and the Square Kilometre Array (SKA; Mellema et al. 2013; Koopmans et al. 2015).

Although these interferometer experiments are yet to detect the 21-cm signal,¹ recent improvements in the flagging of cleaned data

* E-mail: greigh@unimelb.edu.au

¹A detection of excess absorption has been reported by the Experiment to Detect the Global EoR Signature (EDGES; Bowman et al. 2018a) global

and updates to the data analysis and reduction pipelines have yielded the strongest yet upper limits on the 21-cm power spectrum (PS) from LOFAR (Mertens et al. 2020), the MWA (Trott et al. 2020), and HERA (The HERA Collaboration 2022b).

Crucially, 21-cm signal detection is a precision science. The signal is extremely faint, hidden behind astrophysical foregrounds roughly five orders of magnitude brighter than the cosmic signal. Thus, when developing data analysis and reduction pipelines it is vital to test all components that could erroneously inject or remove signal. One vital piece towards this is having sufficiently large simulation volumes that extend well beyond the primary beam and into the side lobes. For example, for an experiment such as the MWA, with a field of view of $\sim 25^2$ deg² at 150 MHz ($z \sim 8.5$), this corresponds to a transverse size of ~ 4 Gpc. Simulations of this scale are currently infeasible.

In this work, we introduce a modified version of the seminumerical simulation code 21CMFAST (Mesinger & Furlanetto 2007; Mesinger, Furlanetto & Cen 2011) specifically tailored towards generating sufficiently large reionization volumes. To achieve this, we make a few simplifying assumptions such as adopting linear structure formation and that the IGM spin temperature is in excess of the CMB temperature ($T_S \gg T_{\text{CMB}}$). Further, we follow a similar approach to Greig, Bolton & Wyithe (2011) and only simulate a restricted volume via considering a frequency-dependent depth.

With these large-volume simulations, redshift-space distortions, signal evolution and sky curvature effects can be correctly simulated to provide realistic 21-cm data sets to apply to data analysis pipelines. The simulations introduced in this work have been added to the software package, WODEN (Line 2022), which produce realistic end-to-end simulations for the MWA Telescope, providing a reference data set for testing the PS amplitude and slope, as well as a comparison against which different analysis techniques can be made (e.g. if the data cleaning and treatment removes sufficient systematic power to reach the cosmological signal).

With such a large simulation volume we can also test and verify several assumptions and approximations that are typically made when being computationally limited to smaller simulation volumes. For example, the impact on the 21-cm PS when simulations are tiled to achieve sufficiently large volumes or whether the 21-cm PS is impacted by large-scale modes longer than the simulation volume. Further, we statistically explore the 21-cm PS by breaking up our large-volume simulation into smaller sub-volumes. We explore the impact of non-Gaussianity in the 21-cm signal on the PS cosmic variance uncertainty following Mondal et al. (2015), Mondal, Bharadwaj & Majumdar (2016), and the potential bias in the 21-cm PS when avoiding the contaminated foreground ‘wedge’ modes relative to the true 21-cm PS from the full simulation as highlighted by Jensen et al. (2016).

The outline of this paper is as follows. In Section 2, we detail our modified version of 21CMFAST and describe the adopted ionizing source prescription. Next, in Section 3 we outline the specifically tailored large-volume simulation for the MWA before investigating various properties and assumptions related to the 21-cm PS. Finally, in Section 4 we conclude with our closing remarks. Unless stated otherwise, quoted quantities are in comoving units and we adopt the cosmological parameters: $(\Omega_\Lambda, \Omega_M, \Omega_b, n, \sigma_8, H_0) = (0.69, 0.31,$

signal experiment, however, its cosmological origins have been disputed in the literature (see e.g. Bowman et al. 2018b; Draine & Miralda-Escudé 2018; Hills et al. 2018; Bradley et al. 2019; Singh & Subrahmanyan 2019; Singh et al. 2022).

0.048, 0.97, 0.81, 68 km s⁻¹ Mpc⁻¹), consistent with recent results from the Planck mission (Planck Collaboration VI 2020).

2 METHOD

In this work, we use a heavily modified version of the seminumerical simulation code 21CMFAST² (Mesinger & Furlanetto 2007; Mesinger et al. 2011). 21CMFAST employs approximate but efficient methods to describe the astrophysics of the EoR which compare favourably to computationally expensive radiative transfer simulations on scales ($\gtrsim 1$ cMpc) relevant to 21-cm interferometer experiments (Zahn et al. 2011). This efficiency is achieved by computing the ionization field using the excursion-set approach (Furlanetto, Zaldarriaga & Hernquist 2004a) which compares the time-integrated number of ionizing photons (following a source prescription) to the number of baryons within spherical regions of decreasing radius, R . Under this method, a simulation voxel at the coordinates (\mathbf{x}, z) is considered ionized if,

$$n_{\text{ion}}(\mathbf{x}, z | R, \delta_R) \geq 1. \quad (1)$$

Here, n_{ion} is the cumulative number of IGM ionizing photons per baryon inside a spherical region of size, R and corresponding smoothed overdensity, δ_R . Simulation cells can additionally be partially ionized, where the ionization fraction is set to $n_{\text{ion}}(\mathbf{x}, z | R_{\text{cell}}, \delta_{R_{\text{cell}}})$, the ionized fraction of the cell smoothed at the minimum smoothing scale, R_{cell} . Explicitly, in this work we ignore the contributions from inhomogeneous recombinations (e.g. Sobacchi & Mesinger 2014) and partial ionizations by X-rays to boost the overall computational efficiency.

For our ionizing sources, we adopt the Park et al. (2019) astrophysical parametrization for high- z galaxies which directly connects the star formation rate (SFR) and ionizing escape fraction to the host dark matter halo mass. The advantage of such an approach is that following some simple conversions, UV luminosity functions can be produced and compared against existing high- z observations. For specific details, we defer the interested reader to the aforementioned work and provide a brief summary of the parametrization below.

First, it is assumed that the typical stellar mass of a galaxy, M_* , can be related to its host halo mass via,

$$M_*(M_h) = f_* \left(\frac{\Omega_b}{\Omega_m} \right) M_h, \quad (2)$$

and that f_* , the fraction of galactic gas in stars, can be expressed as a power law in halo mass with index α_* and this is normalized at a dark matter halo of mass $10^{10} M_\odot$ through $f_{*,10}$,

$$f_* = f_{*,10} \left(\frac{M_h}{10^{10} M_\odot} \right)^{\alpha_*}. \quad (3)$$

Following this, we obtain an estimate for the SFR by dividing the stellar mass by a characteristic time-scale, t_* , corresponding to a fraction of the Hubble time, $H^{-1}(z)$,

$$\dot{M}_*(M_h, z) = \frac{M_*}{t_* H^{-1}(z)}. \quad (4)$$

The escape fraction of UV ionizing photons, f_{esc} , is also assumed to be related to the halo mass by a power law with index α_{esc} and normalization at $10^{10} M_\odot$ via $f_{\text{esc},10}$,

$$f_{\text{esc}} = f_{\text{esc},10} \left(\frac{M_h}{10^{10} M_\odot} \right)^{\alpha_{\text{esc}}}. \quad (5)$$

²In particular, the older C-only version, <https://github.com/andreimesinger/21cmFAST>.

Finally, to account for feedback and inefficient cooling processes which limit small-mass haloes from hosting active, star-forming galaxies a duty-cycle, f_{duty} , is included which describes the fraction, $(1 - f_{\text{duty}})$, of dark matter haloes that cannot host star-forming galaxies,

$$f_{\text{duty}} = \exp\left(-\frac{M_{\text{turn}}}{M_{\text{h}}}\right), \quad (6)$$

with M_{turn} corresponding to this suppression scale. In summary, this ionizing source prescription results in six free parameters: $f_{*,10}$, α_* , $f_{\text{esc},10}$, α_{esc} , t_* , and M_{turn} .

The number of ionizing photons, n_{ion} , inside a spherical region of size, R , is then obtained via

$$n_{\text{ion}} = \bar{\rho}_{\text{b}}^{-1} \int_0^\infty dM_{\text{h}} \frac{dn(M_{\text{h}}, z|R, \delta_R)}{dM_{\text{h}}} f_{\text{duty}} \dot{M}_* f_{\text{esc}} N_{\gamma/\text{b}}, \quad (7)$$

where $\bar{\rho}_{\text{b}}$ is the mean baryon density, $\frac{dn}{dM_{\text{h}}}$ is the conditional halo mass function (HMF), and $N_{\gamma/\text{b}}$ is the number of ionizing photons per stellar baryon. Here, we use the Press–Schechter HMF (Lacey & Cole 1993) normalized to the mean of the Sheth–Tormen HMF (Sheth & Tormen 1999), and $N_{\gamma/\text{b}} = 5000$ representative of a Salpeter initial mass function (Salpeter 1955).

The cosmic 21-cm signal is computed via its brightness temperature contrast against the CMB temperature, T_{CMB} (e.g. Furlanetto et al. 2006):

$$\begin{aligned} \delta T_{\text{b}}(\nu) &= \frac{T_{\text{S}} - T_{\text{CMB}}}{1+z} (1 - e^{-\tau_{\nu_0}}) \\ &\approx 27 x_{\text{H1}} (1 + \delta_{\text{nl}}) \left(\frac{H}{dv_r/dr + H} \right) \left(1 - \frac{T_{\text{CMB}}}{T_{\text{S}}} \right) \\ &\quad \times \left(\frac{1+z}{10} \frac{0.15}{\Omega_{\text{m}} h^2} \right)^{1/2} \left(\frac{\Omega_{\text{b}} h^2}{0.023} \right) \text{ mK}, \end{aligned} \quad (8)$$

where τ_{ν_0} is the optical depth of the 21-cm line, ν_0 , T_{S} is the gas spin temperature, $\delta_{\text{nl}}(x, z)$ is the evolved (Eulerian) overdensity, x_{H1} is the ionization fraction, $H(z)$ is the Hubble parameter, and dv_r/dr is the gradient of the LOS component of the velocity and all quantities are evaluated at redshift $z = \nu_0/\nu - 1$. To aid computational efficiency, in this work we ignore the effects of peculiar velocities (except in Section 3.5) while also assuming that the gas spin temperature is in excess of the CMB temperature ($T_{\text{S}} \gg T_{\text{CMB}}$). Thus, we compute the brightness temperature contrast as

$$\delta T_{\text{b}}(\nu) \approx 27 x_{\text{H1}} (1 + \delta_{\text{nl}}) \left(\frac{1+z}{10} \frac{0.15}{\Omega_{\text{m}} h^2} \right)^{1/2} \left(\frac{\Omega_{\text{b}} h^2}{0.023} \right) \text{ mK}. \quad (9)$$

Typically, 21CMFAST first generates a high-resolution 3D realization of the linear density before applying second-order Lagrange perturbation theory (e.g. Scoccimarro 1998) to obtain the requisite velocity and evolved density fields. It then smooths the evolved fields on to a lower resolution grid to mitigate numerical artefacts. However, to achieve extremely large-volume reionization simulations, we bypass these steps and consider only the linear density field.

Even following this, our largest achievable simulation size is limited to the maximum available memory on our computer network. This limitation is driven by the necessity of performing 3D fast-Fourier transforms (FFTs). To bypass this, we follow the approach of Greig et al. (2011) and take a considerable hit to computational efficiency by storing the 3D data on hard-disk and performing successive 1D and 2D FFTs rather than the expensive 3D FFTs in memory. In effect, this limits our memory requirements to N^2 rather than N^3 and takes advantage of the fact that hard-disk space exceeds total memory (since it is cheaper, though considerably slower to

access). Importantly, by adopting linear theory (initial conditions are generated in Fourier space) and highlighting that the excursion-set algorithm is applied in Fourier space, we can also minimize the number of FFTs between real and Fourier space.

By carefully choosing the order³ of our FFTs, we do not need to calculate the full 3D real-space cubic volume. Instead, we can generate a volume with truncated depth along the line-of-sight equivalent to the desired frequency bandwidth. This results in volumes with much smaller line-of-sight depths than the transverse direction. Importantly, the resultant real-space quantity in the truncated volume is identical to what would be produced if we just performed the full 3D FFT and then only kept the volume to the same desired depth. Thus, the field suffers no loss of power on any scales.

Importantly, although we have made several assumptions to boost computational efficiency at the expense of some physical accuracy, our overall goal is to have a method to generate extremely large 21-cm simulations for testing data analysis and reduction pipelines. For these, physical accuracy is not necessarily the most important requirement, instead, we simply require that the statistics and correlations in our large-volume simulations be representative of those expected of the true signal over an extremely large range of spatial scales and survey footprint. Additionally, for these extremely large-volume simulations we do not include the effects of sky curvature at present. In future, we will look into the impact of sky curvature on large-volume simulations.

3 RESULTS

3.1 Tailored MWA simulation

The MWA upper limits (Trott et al. 2020) were determined from observations spanning 137–197 MHz in two slightly overlapping 30.72 MHz bands. Over these bands, the MWA frequency resolution is ~ 80 kHz, which sets the minimum spatial resolution for our large-volume simulations and the maximal line-of-sight length required (384 channels). At 150 MHz ($z \sim 8.5$), the MWA field of view is $\sim 25^2$ deg², corresponding to a comoving side length of ~ 4 Gpc. To exceed these requirements, we generate a single simulation volume with a comoving transverse side length of 7.5 Gpc over 6400 voxels and 600 voxels (703.1 Mpc) along the line-of-sight direction.⁴ This simulation corresponds to a cell resolution of ~ 1.17 cMpc.⁵ To our knowledge, this is the largest volume 21-cm simulation in the literature.⁶ Using this simulation, we construct two 21-cm light-cones spanning the observing bands ($z = 6.2\text{--}7.5$ and $z = 7.5\text{--}9.4$) by stitching together 2D slices of the 21-cm signal from our simulation volume by linearly interpolating in cosmic time (e.g. Datta et al. 2012, 2014; La Plante et al. 2014; Ghara, Datta & Choudhury 2015; Mondal, Bharadwaj & Datta 2018).

³For example, assuming we want to restrict the line-of-sight (z) depth of our simulation we first 2D FFT along the full $x\text{--}z$ plane (i.e. over the full depth, not the truncated depth) for all y voxels. We then discard all information beyond our desired depth in the z -direction, resulting in our truncated volume. Then, we simply 1D FFT along the y -direction for all remaining (x, z) coordinates to obtain our truncated real-space data.

⁴Such a simulation required ~ 6 d, using 32 CPUs and ~ 10 GB of memory. The majority of this time is spent reading/writing to file, which could be considerably improved using solid-state hard drives.

⁵The raw simulation outputs can be accessed from <https://bradgreig.github.io/large-volume-data/>.

⁶As a verification of our method, we also performed a simulation with a transverse side length of 9.6 Gpc and 8092 voxels.

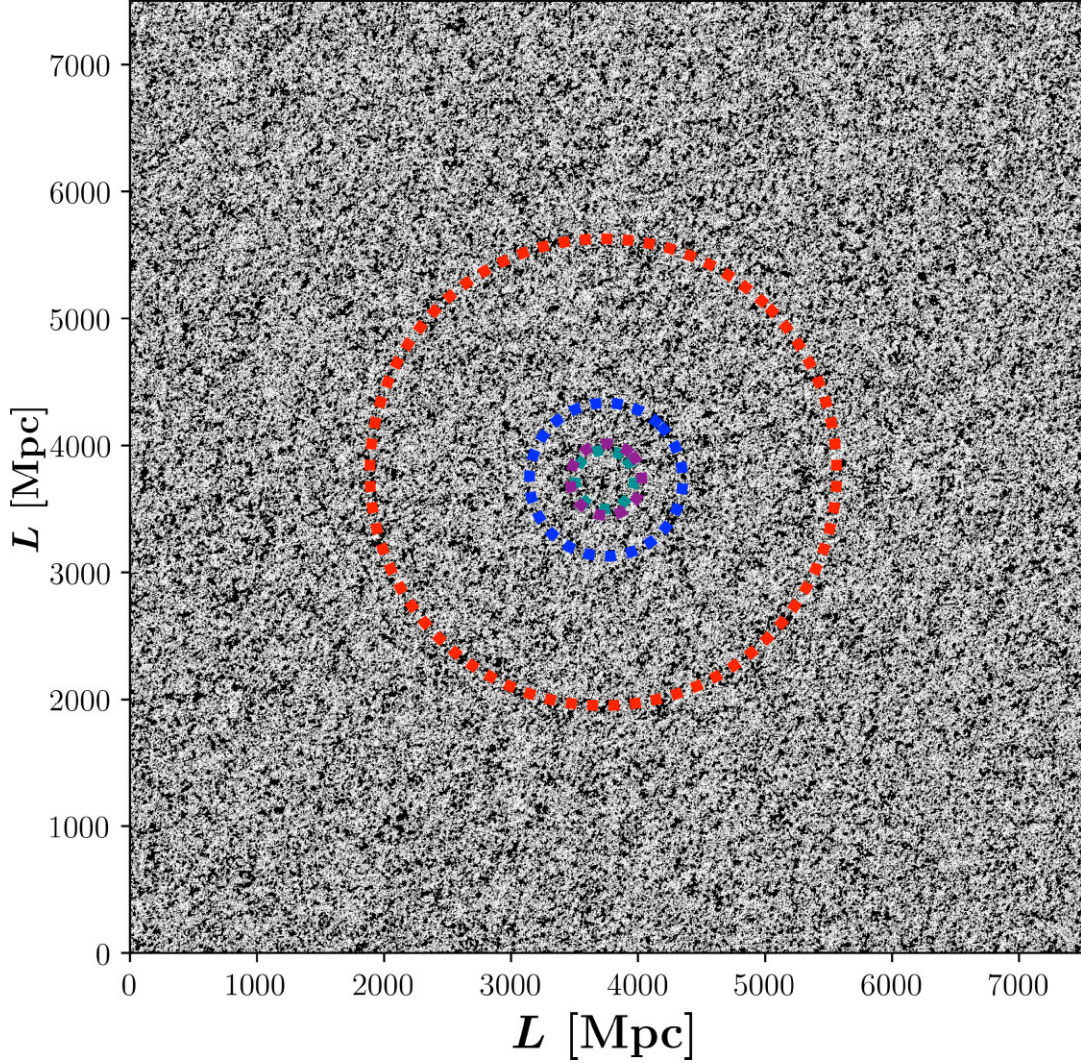


Figure 1. A 2D slice of the ionization field, x_{HI} , obtained from the 7.5 comoving Gpc, 6400 voxel simulation at $z = 8$ (cell resolution ~ 1.17 cMpc). The coloured dashed circles correspond to the field of view for a variety of 21-cm interferometers: MWA (red), HERA (blue), LOFAR (purple), and SKA (teal). See Table 1 for further details.

For this simulation, we adopt $f_{*,10} = 0.05$, $\alpha_* = 0.5$, $f_{\text{esc},10} = 0.08$, $\alpha_{\text{esc}} = -0.5$, $t_* = 0.5$, and $M_{\text{turn}} = 10^{8.7}$, consistent with the fiducial model adopted in Park et al. (2019), which was shown to match all recent observational constraints on the reionization epoch. Additionally, in neglecting inhomogeneous recombinations we are required to set an effective mean free path for the ionizing photons inside the ionized regions, R_{mfp} . We adopt $R_{\text{mfp}} = 15$ comoving Mpc motivated by the fiducial model in Greig & Mesinger (2015).

In Fig. 1, we present a 2D slice of the IGM ionization fraction, x_{HI} , at $z = 8$ for which we obtain a mean IGM neutral fraction of $\bar{x}_{\text{HI}} \sim 0.55$. As a visual demonstration of the physical extent of these simulations, we overlay the respective fields of view for each of the MWA (red), HERA (blue), LOFAR (purple), and SKA (teal). Note that unlike the MWA, LOFAR, and SKA which use phase tracking to follow a single patch of the sky, HERA is a drift-scan experiment (sky rotates over the fixed zenith pointing dishes), thus, here we show its instantaneous field of view. In Table 1, we provide the collecting area of a single station/element and the corresponding field of view for each 21-cm interferometer experiment.

3.2 Exploring large-scale modes

Typically, simulations tailored for a 21-cm interferometer experiment will be designed to match the physical extent of the instruments primary beam. Although problematic for full tests of end-to-end pipelines, for most usage cases this assumption should be sufficient. However, by only simulating a volume equivalent to the primary beam, these simulations forgo modes on much longer scales which would otherwise be present in the observed 21-cm signal. In principle, the absence of these large-scale modes could result in an underestimate of the true large-scale power.

We can test the impact of neglecting modes beyond the physical extent of the primary beam by using our 7.5 Gpc simulation as a reference simulation to represent the true Universe. We then compare the 3D spherically averaged 21-cm PS obtained from our 7.5 Gpc simulation to the PS obtained from the smaller, independent realizations of the 21-cm signal tailored to match each interferometer experiment. Throughout, we define the 21-cm PS as $\Delta_{21}^2(k, z) \equiv k^3 / (2\pi^2 V) \delta \bar{T}_b^2(z) \langle |\delta_{21}(\mathbf{k}, z)|^2 \rangle_k$, where $\delta_{21}(\mathbf{x}, z) \equiv \delta T_b(\mathbf{x}, z) / \delta \bar{T}_b(z) - 1$. Further, when calculating the 21-cm PS we

Table 1. A summary of the collecting area of a single element/station for each of the 21-cm interferometer experiments, the corresponding field of view at 150 MHz and the transverse comoving distance of the simulation required to match the field of view at 150 MHz. Note for HERA, we show the instantaneous field of view as it operates in a drift-scan mode.

Instrument	Collecting area (m ²) (single element)	Field of view (deg ²) (at 150 MHz)	Comoving distance (Mpc)
MWA	21.5	24.7 ²	3977.3
LOFAR	745	3.7 ²	598.8
HERA	154	8.2 ²	1317.3
SKA	1165	3.0 ²	479.0

neglect all $k_{\perp} = 0$ modes. That is, we remove all pure line-of-sight modes as these are not visible to an interferometer⁷ (e.g. Datta et al. 2012).

In Fig. 2, we compare the 21-cm PS obtained from our 7.5 Gpc simulation (black curve) to the mean and 68th percentile scatter obtained from 10 independent realizations of the 21-cm signal. In each panel, we compare the 21-cm PS obtained from the primary field of view of the MWA (red), HERA (blue), LOFAR (purple), and SKA (teal). In all cases, we compute the 21-cm PS over a simulation volume with the same observing bandwidth of ~ 30 MHz (transverse distance is equivalent to the field of view).

Immediately from Fig. 2, we see no obvious deviations at large scales. The mean 21-cm PS determined over the 10 independent realizations for each survey footprint overlaps with the 21-cm PS from the 7.5 Gpc simulation. On the largest scales for each individual instrument, we observe some scatter in the 21-cm PS owing to cosmic variance, where the binned 21-cm power does not have many modes (i.e. Poisson sampling error). Thus, provided the simulation volume exceeds the field of view, this should not be a concern. We can then conclude that omitting large-scale modes beyond the primary field of view does not alter the recovered statistics. This is consistent with the hybrid, large-volume (>1 Gpc) reionization simulations generated by Kim et al. (2016) and the recent study using a reference ~ 1.1 Gpc simulation from 21CMFAST (Kaur, Gillet & Mesinger 2020). In both works, the amplitude of the large-scale 21-cm PS begins to converge for increasing simulation volumes beyond 500 Mpc.

Note, this investigation slightly differs from previous works by Iliiev et al. (2014) and La Plante et al. (2014). Here, in each case the authors explore sub-volumes from their largest reionization simulation, which in the case of Iliiev et al. (2014) is used to determine a minimum volume for reionization simulations to achieve convergence of the large-scale radiative effects. Since this is a sub-division of a larger volume, these smaller simulations contain modes beyond the scale of the sub-volume. In our work, we generate new independent realizations of our smaller volume simulations that do not include modes beyond the scale of the simulation.

3.3 Exploring cosmic variance

The 3D spherically averaged 21-cm PS is calculated by averaging over all Fourier modes that fall within spherical shells. Naturally, the variance of this statistic is then given by the Poisson sampling error

⁷Note, this is somewhat of a crude approximation as interferometers do in fact have a response to these $k_{\perp} = 0$ modes, although it is suppressed due to the tapered response. While this does not impact the results of this work, realistically one should include the true instrument response rather than taking this approximation when comparing simulations to observations.

based on the total number of Fourier modes in each spherical shell,

$$\sigma_{\Delta_{21}^2}(k, z) = \Delta_{21}^2(k, z) \sqrt{\frac{(2\pi)^2}{Vk^2\Delta k}}. \quad (10)$$

Here, $\Delta_{21}^2(k, z)$ is the dimensional 21-cm PS measured within a volume, V , with spherical shells separated by Δk .

However, this assumes that the 21-cm PS fully encodes all the information about the 21-cm signal (i.e. that the 21-cm signal is Gaussian). Instead it is well known that the 21-cm signal is highly non-Gaussian (e.g. Furlanetto et al. 2004a; Furlanetto, Zaldarriaga & Hernquist 2004b; Morales & Hewitt 2004; Cooray 2005), in which case the higher order moments of the 21-cm signal will be non-zero. As a result of this, Mondal et al. (2016) have shown that the cosmic variance uncertainty on the 21-cm PS is instead,

$$\sigma_{\Delta_{21}^2}(k, z) = \Delta_{21}^2(k, z) \sqrt{\frac{(2\pi)^2}{Vk^2\Delta k} + \frac{T(k, k)}{V}}, \quad (11)$$

where $T(k, k)$ is the Trispectrum component which arises from the four-point correlation function of the brightness temperature fluctuations.

To explore this non-Gaussianity in the 21-cm PS cosmic variance, we extract a large number of sub-volumes from our large 7.5 Gpc simulation. In particular, we extract sub-volumes equivalent in spatial extent to the field of view for each 21-cm interferometer experiment (as listed in Table 1), each with the same 30 MHz bandwidth along the line of sight. Note, here we are specifically interested in the sampling statistics of the 21-cm PS and thus we require large numbers of volumes. Since our fiducial simulation volume is larger than any instrument field of view it is computationally more convenient to simply extract sub-volumes rather than to generate new independent realizations. Additionally, since these sub-volumes are created from our large-volume simulation they then have the correct large-scale power.

To create these sub-volumes, we first split the full 7.5×7.5 Gpc² sky-area into our requisite smaller volumes without overlapping regions. Then, to increase our number statistics, we generate an inset sky-area offset from the simulation edge by half the size of the respective instrument field of view. We then additionally split this inset area into non-overlapping sub-volumes. In doing so, we obtain a total of 5, 41, 221, and 365 different sub-volumes from our single, large-volume simulation for the MWA, HERA, LOFAR, and the SKA, respectively. By considering sub-volumes equivalent to each instrument's field of view, we consider volumes $\sim 2-10 \times$ larger than those considered by Mondal et al. (2016) and Mondal, Bharadwaj & Majumdar (2017). Further, by splitting up our large-volume simulation we can obtain up to ~ 7 times as many realizations, increasing our statistical sampling of the numerical cosmic variance uncertainty, albeit at the expense of notably less physics rich simulations.

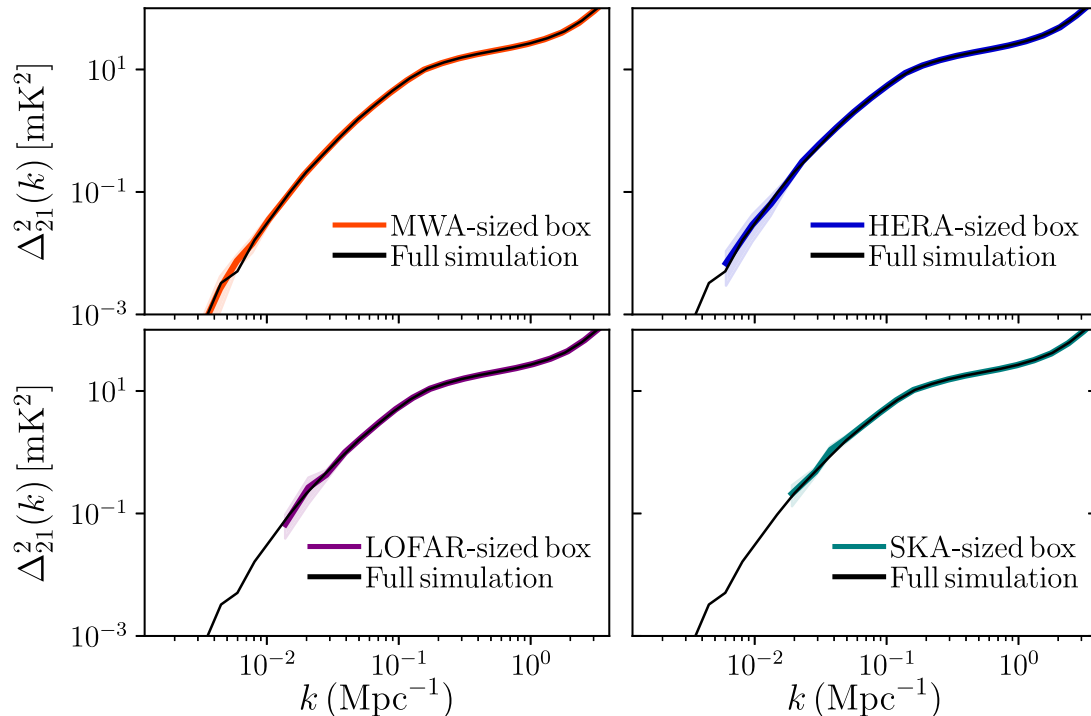


Figure 2. The 21-cm PS from simulations tailored to the field of view of specific 21-cm interferometers. The black curve corresponds to the 21-cm PS from the full $7500 \text{ Mpc} \times 7500 \text{ Mpc} \times 500 \text{ Mpc}$ simulation. The red, blue, purple, and teal curves correspond to simulation volumes with angular extents equivalent to the field of view of the MWA, HERA, LOFAR, and the SKA, respectively. The shaded region corresponds to the 1σ scatter from 10 realizations with different initial conditions.

In Fig. 3, we show the resultant 1σ uncertainty on the 21-cm PS obtained from our distribution of sub-volumes (solid curves) for each 21-cm interferometer experiment. In contrast, we also show the expected theoretical estimate of the uncertainty from equation (10) if we assumed Gaussian statistics (dashed curves). The dotted curve corresponds to the non-Gaussian (Trispectrum) component of the cosmic variance as highlighted in equation (11).

On relatively large scales ($k < 0.5 \text{ Mpc}^{-1}$), we find that the true cosmic variance uncertainty is well approximated by a Gaussian component for the theoretically expected error. However, progressing to intermediate and smaller scales, the amplitude of the non-Gaussianity increases significantly, with the Trispectrum component dominating by $\geq 10\times$ that of the Gaussian term. This transition scale of $k > 0.5 \text{ Mpc}^{-1}$ for a non-Gaussian dominated cosmic variance uncertainty for the 21-cm PS is consistent with that of Mondal et al. (2016, 2017) for a similar stage of reionization ($\bar{x}_{\text{H}1} \sim 0.55$). On these scales, as the EoR progresses, the growth and percolation of the ionized regions drive the non-Gaussianity in the cosmic 21-cm signal. While not shown here, as we progress towards the end of reionization the non-Gaussian signal continues to increase (e.g. Mondal et al. 2017; Majumdar et al. 2018; Hutter et al. 2020). Thus, the amplitude of this Trispectrum component will continue to increase in amplitude and as a result dominate over the Gaussian component over most if not all scales. Importantly, we find the relative amplitude (the amplitude simply scales as $\propto V^{-1}$) and shape of the full cosmic variance uncertainty to be consistent across the differently sized sub-volumes. This indicates no additional sources of cosmic variance uncertainty due to the finite-size of the simulation volume considered.

Note however, that while the non-Gaussian term for the 21-cm PS cosmic variance begins to dominate on scales larger than $k >$

0.5 Mpc^{-1} , this need not be too concerning for 21-cm experiments. Typically, on these scales, the 21-cm PS will be dominated by the instrument thermal noise (see fig. 2 of Greig, Mesinger & Koopmans 2020) making these two terms comparable. Nevertheless, care must be taken when estimating the 21-cm PS cosmic variance error.

One thing we have not considered is that the cosmology of our sub-volumes may not be the same as our single, large-volume simulation. That is, the resultant cosmology of the sub-volumes may not be flat like our full simulation. This was explored in the context of the position-dependent bispectrum (Giri et al. 2019). However, for this work, since we are only interested in the statistical uncertainty over a large number of realizations, not the amplitude of the 21-cm PS itself this likely has little bearing on our results. If anything, it should be encapsulated by the non-Gaussian (Trispectrum) component. We will investigate this in the future.

3.4 Impact of foreground ‘wedge’ avoidance on cosmic variance

The cosmic 21-cm signal observed by radio interferometer experiments is measured in visibility space (uv coverage), which is the 2D Fourier transform of the signal in the image (sky) plane. Measuring the PS then requires gridding these visibilities. However, these visibilities are frequency dependent. Line-of-sight power in Fourier space (frequency dependent) can then leak into the transverse (frequency independent) Fourier modes. This results in the well-known ‘wedge’ feature in cylindrical 2D Fourier space, where measured power within this regime is contaminated (Datta, Bowman & Carilli 2010; Morales et al. 2012; Parsons et al. 2012; Vedantham, Shankar & Subrahmanyam 2012; Trott, Wayth & Tingay 2012; Thyagarajan et al. 2013; Liu, Parsons & Trott 2014a, b; Thyagarajan et al. 2015a, b; Pober et al. 2016; Murray & Trott 2018). While it is plausible to

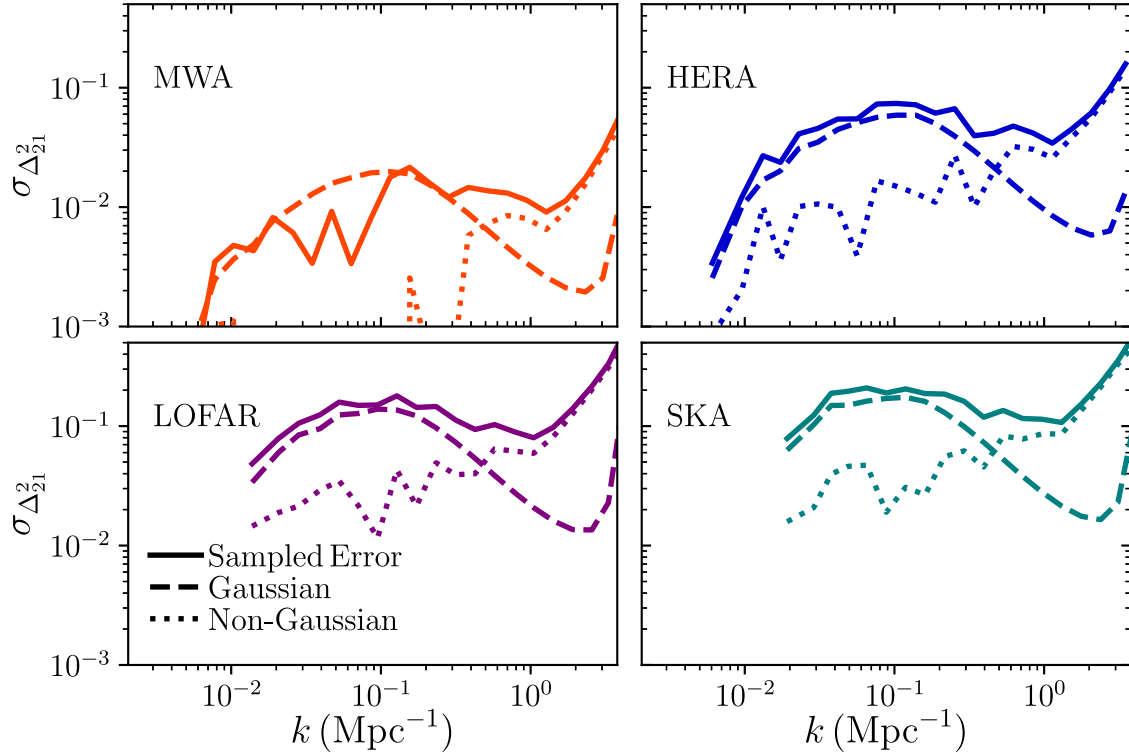


Figure 3. The 1σ variation in the 21-cm PS estimated from sub-volumes of size equivalent to the 21-cm experiment field of view (solid curves) compared to the theoretically expected variation assuming Gaussian errors (dashed) and the non-Gaussian contribution to the sample variance (dotted). *Top left: MWA, top right: HERA, bottom left: LOFAR, and bottom right: SKA.*

mitigate or ‘clean’ these contaminated modes (see e.g. Chapman & Jelić 2019 for a comprehensive review), it is typically easier to simply avoid this contaminated region entirely, measuring the PS in the relatively pristine region above the ‘wedge’.

The spatial location of this ‘wedge’ in 2D cylindrical space is determined by

$$k_{\parallel} = mk_{\perp} + b, \quad (12)$$

where k_{\parallel} and k_{\perp} are the line-of-sight and transverse Fourier modes, respectively. The constant b corresponds to an additive buffer region above the ‘wedge’ of $\Delta k_{\parallel} = 0.1 h \text{ Mpc}^{-1}$ extending beyond the horizon limit while,

$$m = \frac{D_C H_0 E(z) \sin(\theta)}{c(1+z)}. \quad (13)$$

Here, D_C is the comoving distance, H_0 is the Hubble constant, $E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$, and θ is the angular radius of the field of view which we conservatively take as $\theta = \pi/2$ (i.e. observing down to the horizon).

Previously, we numerically estimated the cosmic variance in the 21-cm PS measured over the full 30 MHz bandwidth simulation volume for each interferometer experiment. To highlight the impact of ‘wedge’ avoidance on estimating the 21-cm PS, in particular on the amplitude of the statistical uncertainty, we perform the same analysis as in Section 3.3 instead calculating the 21-cm PS using only Fourier modes above the ‘wedge’.

In Fig. 4, we compare the 1σ cosmic variance uncertainty after ‘wedge’ avoidance (dashed curves) relative to the 21-cm PS estimates from the full simulation volume (solid curves). As expected, in avoiding the ‘wedge’ region, we limit the total number of Fourier modes available in each spherical shell to measure the 21-cm PS.

Thus, quite simply, our Poisson uncertainty increases owing to the smaller number of available modes driving the increasing uncertainty error. Further, as the number of available modes quickly diminishes as we move to larger spatial scales (smaller k) the cosmic variance uncertainty rapidly increases. For smaller scales, the cosmic variance uncertainty still becomes dominated by the Trispectrum component (as indicated by similar upticks in amplitude at $k > 1 \text{ Mpc}^{-1}$ as highlighted in Fig. 3), though on slightly smaller scales than found previously. Nevertheless, the non-Gaussianity in the cosmic variance uncertainty is still important when performing ‘wedge’ avoidance.

Importantly, the differences between the two cases are amplified by our survey geometry. In general, our simulation volumes are considerably larger along the transverse direction relative to the line of sight. Thus, previously, when measuring the 21-cm PS over the full volume, our number statistics for the k_{\perp} modes were considerably larger than for the k_{\parallel} modes. Now, when measuring the 21-cm PS using foreground avoidance, most of these k_{\perp} modes are excluded since they fall within the ‘wedge’. Nevertheless, despite this we still recover a notably lower cosmic variance uncertainty with the MWA relative to the other instruments due to the larger field of view. The larger field of view ensures better numerical statistics for all sampled Fourier scales, driving the Poisson uncertainty lower.

In following this avoidance strategy, we also significantly limit the largest scale modes accessible to our 21-cm PS measurement. In all cases, the 21-cm PS is only measurable down to $k \sim 0.1 \text{ Mpc}^{-1}$, with no measurable modes beyond this limit. This limit is simply set by the adopted size of our buffer ($b = 0.1 h \text{ Mpc}^{-1}$) from equation (12) which is somewhat arbitrary. In principle, it can be reduced by considering either larger bandwidths or with an improved understanding of our systematics. Note however, that the smallest accessible scale for each instrument is not exactly $k \sim 0.1 \text{ Mpc}^{-1}$,

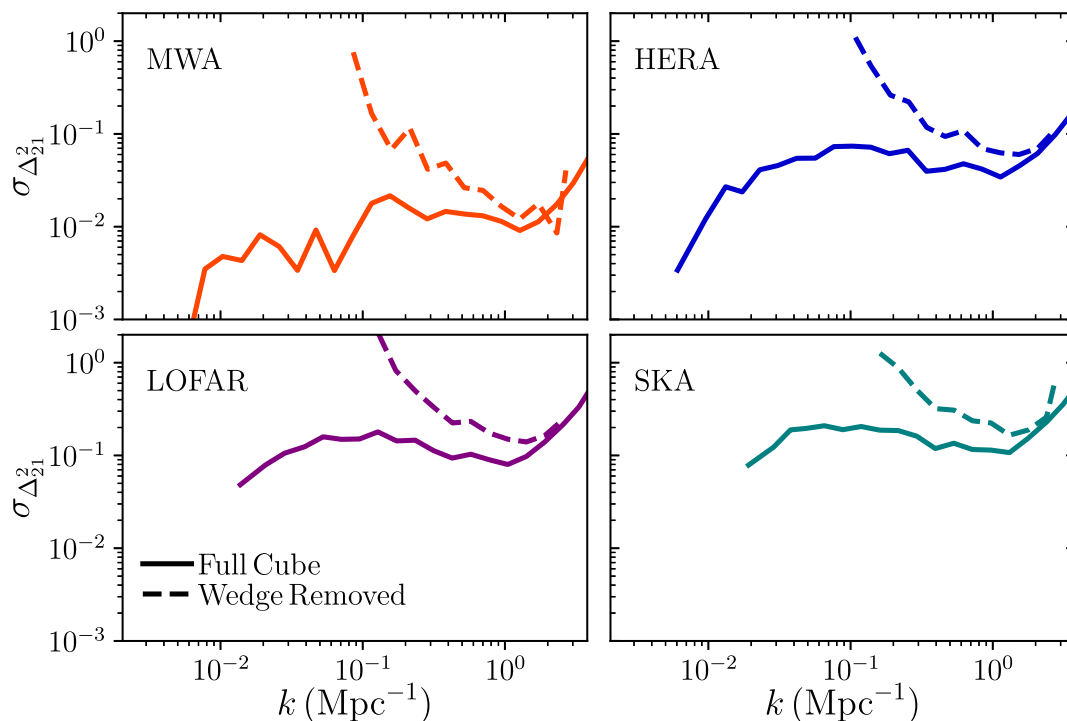


Figure 4. The 1σ variation in the 21-cm PS estimated from sub-volumes of size equivalent to the 21-cm experiment field of view when using the full simulation box (solid curves) and after removing the foreground contaminated ‘wedge’ modes (dashed). *Top left:* MWA, *top right:* HERA, *bottom left:* LOFAR, and *bottom right:* SKA.

but rather slightly larger. As we additionally ignore all $k_{\perp} = 0$ modes, from equation (12) the minimum scale is simply b plus m times the smallest non-zero k_{\perp} in our simulation volume. Thus, the minimum k for the MWA is smaller than the SKA due to the smaller k_{\perp} (larger transverse scale). Note, although this $k \sim 0.1 \text{ Mpc}^{-1}$ limit appears restrictive, the 21-cm PS becomes less sensitive to the astrophysics of reionization below this limit (e.g. Greig & Mesinger 2015; Cohen, Fialkov & Barkana 2018). Thus, we do not particularly lose much constraining power for distinguishing astrophysical models by not being sensitive below these scales.

In Fig. 5, we present the fractional increase in cosmic variance when removing the foreground contaminated wedge modes. We find that the relative increase is comparable across all scales irrespective of the total simulation volume, highlighting that the increase in the cosmic variance is primarily driven by the fraction of excised Fourier modes (i.e. increased Poisson noise in each bin due to having fewer modes free of foregrounds).

3.5 Biased estimates of the 21-cm PS following ‘wedge’ avoidance

Previously, we explored the impact of foreground avoidance on the resultant cosmic variance uncertainty on a 21-cm PS measurement. Next, we consider whether the 21-cm PS measured following foreground avoidance can be biased relative to the 21-cm PS measured from a full simulation volume. This is particularly pertinent given that recent upper limits on the 21-cm PS measured by LOFAR (Mertens et al. 2020), MWA (Trott et al. 2020), and HERA (The HERA Collaboration 2022b) have for the first time allowed Bayesian astrophysical parameter inference to be employed to understand the physics of reionization (Ghara et al. 2020, 2021; Mondal et al. 2020; Greig et al. 2021a, b; The HERA Collaboration

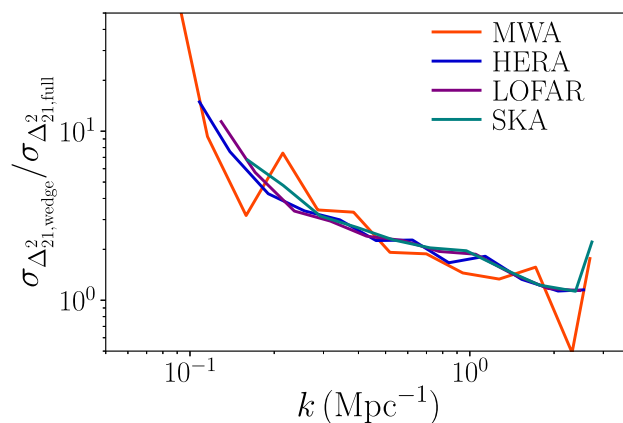


Figure 5. The fractional increase in the cosmic variance on the 21-cm PS when removing the foreground wedge contaminated modes ($\sigma_{\Delta_{21,\text{wedge}}}^2$) compared to the cosmic variance estimated from the full 21-cm simulation ($\sigma_{\Delta_{21,\text{full}}}^2$).

2022a). However, when these inference pipelines are applied to these upper limits, the simulated 21-cm PS is generated from the full simulation volume, rather than the characteristics of the observational data. This is in particular pertinent for the MWA and HERA analyses where foreground avoidance is adopted in the data analysis.

While likely irrelevant for the present upper limits since these are not very aggressive and only disfavour relatively extreme models of reionization. Nevertheless, as we approach a first detection, these inference pipelines will be vital to building our understanding of the

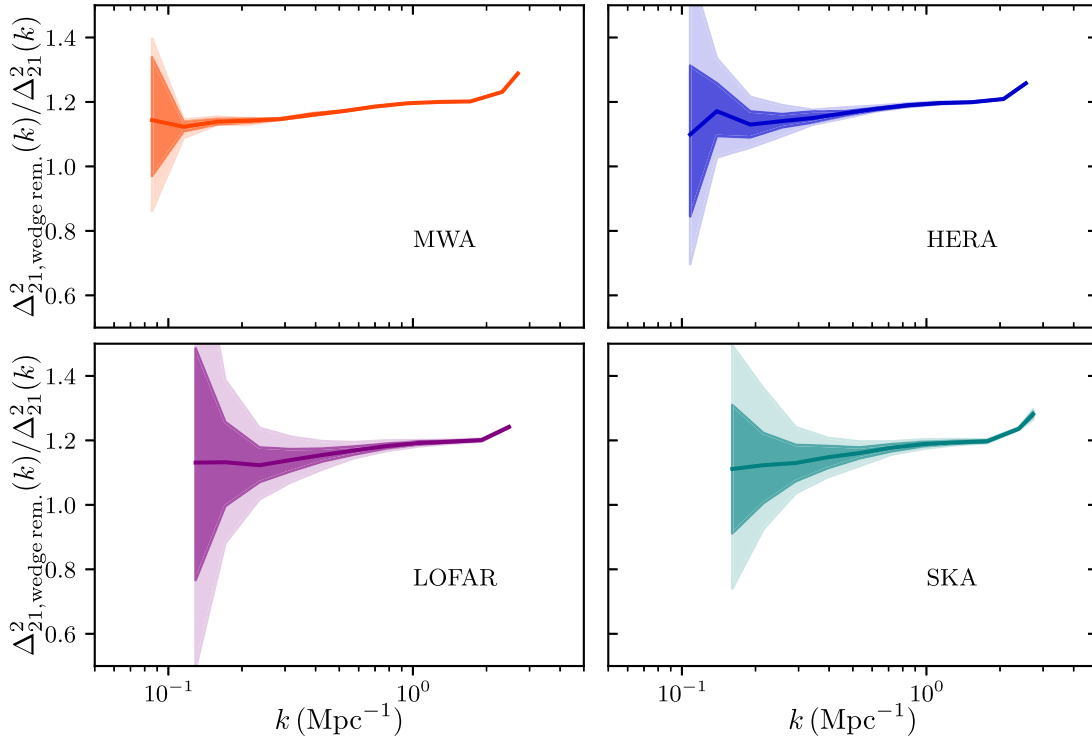


Figure 6. The ratio of the 21-cm PS with wedge modes removed (foreground avoidance) relative to the 21-cm PS from the full simulation volume. Solid curves correspond to the mean ratio, while dark and light shaded regions correspond to the 68th and 95th percentiles, respectively. *Top left:* MWA, *top right:* HERA, *bottom left:* LOFAR, and *bottom right:* SKA.

galactic physics driving reionization. Thus, it is important to explore the relative impact of any biases on the modelled 21-cm PS when not accounting for the characteristics of the observation.

First, the observed 21-cm PS is asymmetric owing to redshift-space distortions from the peculiar motions of the gas along the line of sight. Thus unlike in the earlier sections, we must include the effects of redshift-space distortions on our estimate of the 21-cm brightness temperature signal (note, the inference pipelines above include these effects). These redshift-space distortions increase the relative power along the line-of-sight direction, which could amplify any biases when measuring the 21-cm PS following foreground avoidance. This is because modes are preferentially excited transverse to the line of sight, resulting in a higher amplitude signal after binning over all possible modes within spherical shells. Thus here, we compute the 21-cm brightness temperature signal for the full simulation following the multiplication of equation (9) by the pre-factor $(1 + \frac{dv_r}{Hdr})^{-1}$ and computing the anisotropic 21-cm PS.

In Fig. 6, we present the mean (solid curve) and the 68th and 95th percentiles (dark and light, respectively) for the ratio of the 21-cm PS with the ‘wedge’ removed (avoidance) relative to the 21-cm PS from the full simulation volume. Here, these are determined using the same sub-volumes as obtained previously for each individual experiment from the full 7.5 Gpc simulation. It is clearly evident that for all survey areas, the 21-cm PS measured following foreground avoidance returns a biased estimate of the true 21-cm PS obtained from the full simulation volume. On average, we find the relative amplitude of this bias to be relatively modest, at ~ 10 – 20 per cent over all measured k -scales. On the largest scales (small k), when foreground avoidance has limited sampling of the Fourier modes (see Fig. 4) this can result

in increases/decreases in the 21-cm PS amplitude of larger than 40 per cent.

Although fairly modest in amplitude, it is important to note that these results are model dependent. That is, this bias corresponds specifically to our particular source parametrization and frequency band ($z \sim 7.5$ – 9.4). Therefore, it is plausible that the amplitude of this bias could increase under different astrophysical models or stages of reionization. Hence, this could have important consequences for parameter inference studies if the statistics of the simulated models are not consistent with the characteristics of the observation, potentially biasing the inferred astrophysical parameters. Although the model dependence of the work is important to consider here, in previous sections it is less relevant. When investigating the cosmic variance uncertainty the relative amplitude of the uncertainty is dependent on the number of modes within each spherical shell when estimating the 21-cm PS, not the 21-cm PS amplitude itself.

To more clearly illustrate the relative increase in amplitude in the 21-cm PS following foreground avoidance, in Fig. 7 we show in the same panel the mean ratio of the 21-cm PS with/without foreground avoidance for all interferometer experiments. This clearly demonstrates that the relative increase in the 21-cm PS amplitude following foreground avoidance is consistent at 10–20 per cent across all the different survey footprints.

This 10–20 per cent bias is broadly consistent with a similar investigation performed by Jensen et al. (2016). Here, this bias was determined for two astrophysical models over the full reionization history. For their fiducial model, which more closely matches our model and at a similar stage of reionization, $\bar{x}_{\text{H I}} \sim 0.5$, they find a similar bias of 10–20 per cent, however, as a decrement rather than an amplification. Likely, these differences arise due to the specific modelling of the ionizing sources, but more likely due to the

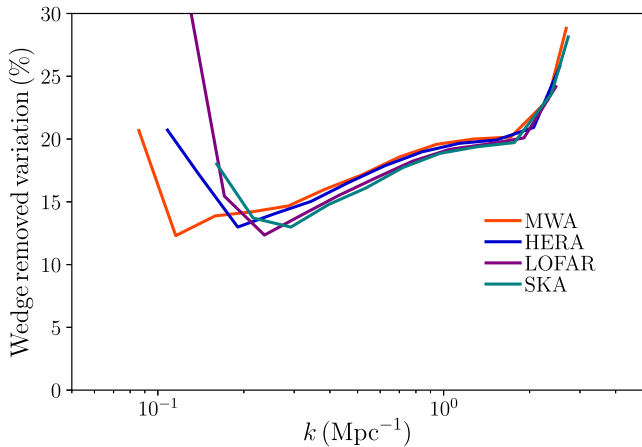


Figure 7. The mean ratio in the 21-cm PS with wedge modes removed compared to the 21-cm PS from the full simulation volume. The red, blue, purple, and teal curves correspond to simulation boxes with fields of view equivalent to the MWA, HERA, LOFAR, and the SKA, respectively.

simplifications present in our modelling (e.g. linear theory) relative to their N -body simulations. Importantly, though, they only consider a single realization for this exploration. For sub-volumes equivalent in size to their single simulation (corresponding to the LOFAR field of view) in the lower left panel of Fig. 6 we can accommodate individual models with a negative bias. Factoring in the modelling differences mentioned above, we are confident that these two works are broadly consistent.

Note, there also appears to be a slight gradient in the relative bias for the different experiments. The SKA, with the smallest survey footprint has the lowest amplitude whereas the MWA, with the largest footprint has the highest amplitude. However, this is simply due to the geometry of the survey volume. Foreground avoidance with the MWA removes on average more Fourier modes in each spherical shell below the ‘wedge’ than that of the SKA relative to the 21-cm PS computed from the full simulation volume. This is because the MWA has a considerably larger transverse size (higher sampling of transverse modes) which results in a higher fraction of modes cut at the same fixed k than the SKA. Equally, the uptick at the largest scales occurs earliest for the SKA compared to the other, larger area surveys. This occurs for the same reason as discussed earlier, the removal of the $k_{\perp} = 0$ modes coupled with the k_{\parallel} needing to satisfy equation (12).

3.6 Impact of tiling reionization simulations

Typically, when it is infeasible to generate a single, large-volume simulation, the requisite size and volume can be achieved by stitching together many smaller simulations generated using periodic boxes (see e.g. Nasirudin et al. 2020). Although sufficient volumes are achieved in this manner, the statistics of the simulations will only be accurate on the scales modelled by the smaller simulation. That is, it is impossible to inject information on scales beyond the largest scale of the smallest simulation. Using our 7.5 Gpc simulation, we can explore the relative impact of tiling simulations on the large-scale information in the 21-cm PS.

The tiling of periodic simulations additionally leads to repeating structures on scales characteristic of the simulation volume. When considering discrete objects such as galaxies in a redshift survey, these repeating structures can be minimized through rotating the smaller simulations prior to tiling. Unfortunately, since the 21-cm

signal arises from the neutral gas, it is a continuous rather than discrete quantity preventing the usage of rotations. Rotations would lead to edge effects and disjoint ionized regions on the boundaries of the smaller simulations leading to strange artefacts in the statistics. Thus, for 21-cm studies, we can only consider tiled simulations.

To investigate the impact on the large-scale power due to tiling, we consider three smaller volumes and tile each to match our single 7.5 Gpc, 6400 voxel simulation volume. We generate 50 independent realizations of the smaller simulations with 200, 400, and 600 voxels per side length, corresponding to lengths of ~ 234.4 , 468.8, and 703.1 Mpc. We tile the same individual simulation to the large-volume footprint, thus we obtain 50 independent large-volume tiled simulations. For each, we generate the full 21-cm light-cone and only consider a line-of-sight depth of 30 MHz, consistent with our large-volume simulation.

In Fig. 8, we compare the 21-cm PS obtained from our 7.5 Gpc simulation (black curve) to the 21-cm PS obtained from an equivalent tiled 7.5 Gpc simulation using the 200 (red), 400 (blue), and 600 (purple) voxel smaller simulations. In each, the vertical dashed line corresponds to the largest scale (fundamental frequency) sampled by the smaller simulations ($k_f = 2\pi/L_{\text{small}}$). The shaded region corresponds to the 68th percentile scatter obtained from our 50 independent realizations. As expected, the 21-cm power drops away on scales below the fundamental frequency. This highlights that information on modes not originally simulated by the smaller volume simulations cannot be artificially created to match the full large-volume simulation. Thus, a tiled simulation volume has a limited regime of validity restricted by the scale of the smaller simulations.

On scales just above the fundamental frequency, we find the 21-cm power in the tiled box underestimates the power from the full simulation. In some instances this can be as much as a factor of ~ 2 – 3 . Although the physical scale of these spherically averaged Fourier modes are sampled by the original smaller volume simulation, on these scales we are still missing the contribution from Fourier modes which contain at least one-component mode longer than the small-volume simulation (e.g. k_x , k_y , or $k_z < k_f$). These individual modes contain no power, but contribute to the total number of modes within the bin, resulting in an overall reduction to the averaged power in that bin. Thus, the 21-cm power from the tiled box will underestimate the full box power on all scales until the fraction of Fourier modes with individual components larger than the smaller box is sufficiently small. Therefore, care must be taken to adequately produce sufficiently sized smaller volumes prior to tiling to minimize the loss of power on the largest scales.

In addition to underestimating the 21-cm power following the tiling of smaller periodic simulations, it is also important to note that we will underestimate the cosmic variance uncertainty in our tiled simulation volume. Since in the act of tiling we produce repeated copies of the spatial information from our smaller simulations we cannot artificially increase the variance within a sampled Fourier bin. That is, the measured sample variance is fixed to the variance (i.e. Poisson fluctuations) measured from the smaller volume simulation. However, since we have scaled up our total volume, following from equation (11) the sample variance decreases as $\propto 1/V$. Thus our estimated cosmic variance from our tiled, large-volume simulation must be scaled by the increased volume. In effect, the tiled simulation underestimates the cosmic variance error by $\propto \sqrt{N_{\perp}^2 \times N_{\parallel}}$, where N_{\perp} is the number of copies required to achieve the desired transverse scale and N_{\parallel} is the number of copies required along the line of sight.

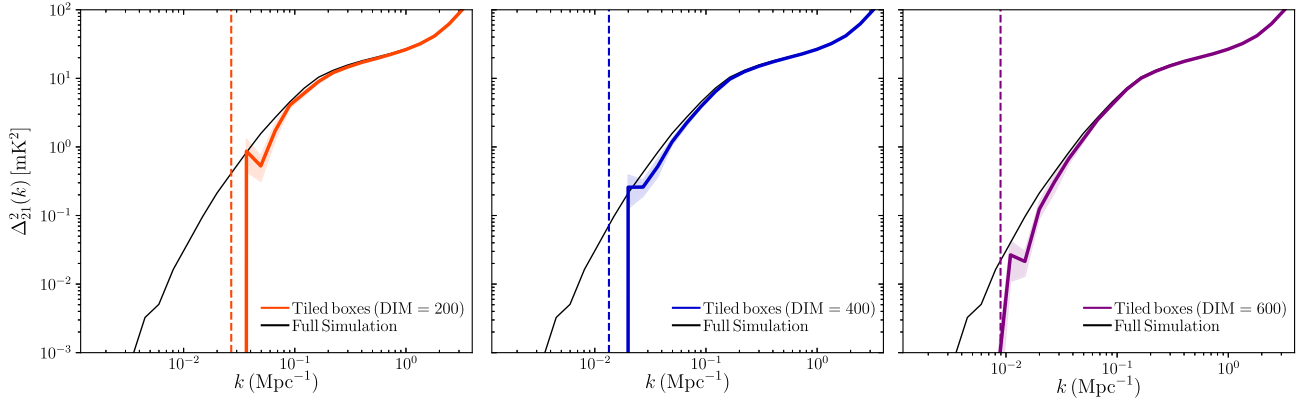


Figure 8. The 21-cm PS obtained after tiling smaller volume simulations. The black curve corresponds to the 21-cm PS of the full $7500 \text{ Mpc} \times 7500 \text{ Mpc} \times 500 \text{ Mpc}$ simulation, whereas the red, blue, and purple curves correspond to the 21-cm PS obtained from an equivalent simulation volume after tiling a single simulation box containing 200, 400, and 600 voxels per side length, respectively. Vertical dashed lines correspond to the fundamental frequency, $2\pi/L_{\text{small}}$, for each small volume.

In Fig. 9, we highlight this scaling by showing the ratio of the cosmic variance uncertainty from the full simulated volume compared to that obtained from tiling the smaller volume simulations to achieve the same volume. In each, the tiled simulation cosmic variance underestimates the true cosmic variance, as indicated by the ratio sitting well above unity. The dashed horizontal lines correspond to the expected scaling ($\propto \sqrt{N_{\perp}^2 \times N_{\parallel}}$) by which we underestimate the true cosmic variance in each tiled simulation volume. That is, the cosmic variance in the tiled simulation must be decreased by this amount to match the true cosmic variance from the full simulation volume. For larger k (i.e. $k > 0.5 \text{ Mpc}^{-1}$), the impact of the non-Gaussian (Trispectrum) contribution to the cosmic variance becomes increasingly important, modifying this simple scaling relation.

Importantly, here we have only considered the power extracted from the full simulation volume. However, if one also considers the instrumental response of the beam (i.e. beam multiplied by the tiled signal), then additional artefacts in the estimated power could be expected due to mode mixing and/or power leaking in from the side lobes. We leave an investigation into these effects to future work.

4 CONCLUSION

In preparation for the wealth of data expected from large-scale interferometer experiments such as the MWA, LOFAR, HERA, and the SKA, we must rigorously test and validate the full data analysis and reduction pipelines. For experiments such as the MWA, with a field of view of $\sim 25^2 \text{ deg}^2$ at 150 MHz ($z \sim 8.5$) this requires reionization simulations in excess of $\sim 4 \text{ Gpc}$. To explore the effects of side lobes away from the primary beam and the potential for large-scale power leakage, even larger volumes are required.

We introduced a modified version of the seminumerical reionization simulation code 21CMFAST (Mesinger & Furlanetto 2007; Mesinger et al. 2011) specifically tailored towards generating these extremely large-volume simulations. To achieve this, we limit the line-of-sight depth of the simulations to specific observing bandwidths (though fully sample the large-scale modes), allowing the transverse scales to be extremely large. Further, we forego some level of accuracy by limiting structure formation to linear theory. However, for the purposes of testing and validation, provided the simulated 21-cm signal behaves in a similar manner (i.e. correct statistics and correlated behaviour) to the expected signal, these assumptions are

sufficient. Using this, we generated a 30 MHz bandwidth ($\sim 500 \text{ Mpc}$) simulation with 6400 voxels over a 7.5 Gpc transverse side length ($\sim 1.17 \text{ cMpc resolution}$). This simulation was specifically tailored for the recent 2020 MWA upper limits (Trott et al. 2020).

With such a large-volume simulation in hand, we then explored the statistical behaviour of the 21-cm PS and assumptions made in the literature when computational limitations restrict the maximum achievable simulation volumes. We performed these using simulation volumes equivalent in size to the field of view for each of the 21-cm interferometer experiments, MWA, LOFAR, HERA, and the SKA. We found:

(i) Finite simulation volumes do not exhibit any loss or increase in large-scale power from missing modes longer than the finite simulation volume.

(ii) The cosmic variance uncertainty on a 21-cm PS measurement at $k \gtrsim 0.5 \text{ Mpc}^{-1}$ is dominated by the non-Gaussianity in the 21-cm signal by $\gtrsim 10\times$ over that expected under the typical Gaussian assumption. This is consistent with that found earlier by Mondal et al. (2016) at a similar stage of reionization ($\bar{x}_{\text{H I}} \sim 0.55$).

(iii) The excision of contaminated foreground wedge modes significantly increases the amplitude of cosmic variance due to the reduction in the number of available Fourier modes. Further, it restricts the largest scale modes accessible.

(iv) Measuring the 21-cm PS following wedge mode excision results in a modestly biased estimate (increase of $\sim 10\text{--}20$ per cent) of the true underlying 21-cm PS. Though, the relative amplitude will be heavily model dependent. This bias is broadly consistent with that found by Jensen et al. (2016) and should be an important consideration for reionization astrophysical parameter inference studies.

(v) Tiling smaller simulations to achieve extremely large volumes results in an underestimate of the 21-cm power on scales smaller than the small box fundamental frequency $k_{\text{f}} \gtrsim 2\pi/L_{\text{small}}$. Therefore, care must be taken when constructing smaller volume simulations for tiling to ensure the relevant scales of interest are sufficiently sampled.

(vi) Tiling smaller simulations additionally results in underestimating the true cosmic variance of the desired large-volume simulation. The underestimation can be compensated for by scaling the uncertainty by $\propto \sqrt{N_{\perp}^2 \times N_{\parallel}}$ which accounts for the number of

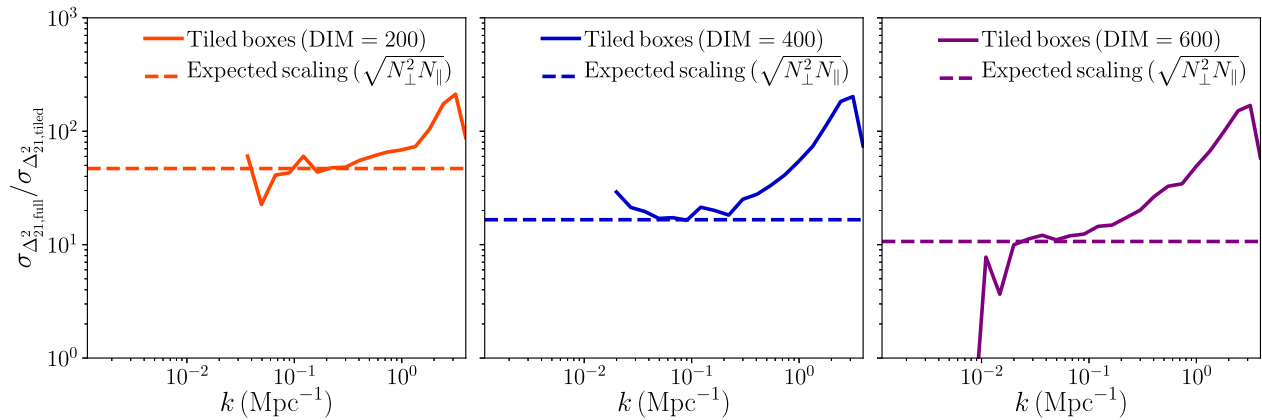


Figure 9. The ratio of the cosmic variance estimated from the full 7500 Mpc \times 7500 Mpc \times 500 Mpc simulation to the cosmic variance estimated from tiling smaller volume simulations to achieve comparable volumes. Again, we consider tiling a single simulation box containing 200 (red), 400 (blue), and 600 (purple) voxels per side length. The cosmic variance of the tiled simulations underestimates the true cosmic variance by an amount proportional to the total number of copies required to be tiled ($\propto \sqrt{N_{\perp}^2 \times N_{\parallel}}$) as estimated by the horizontal dashed curves (see the text for further details).

repeated copies of the small-volume simulation required to match the transverse and line-of-sight scales, respectively.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their insightful comments. We also thank Nichole Barry, Sambit Giri, Garrelt Mellema, Andrei Mesinger, Rajesh Mondal, and Rachel Webster for useful discussions relating to this work. Parts of this research were supported by the ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. CMT is supported by an ARC Future Fellowship under grant FT180100321. Parts of this work were performed on the OzSTAR national facility at Swinburne University of Technology. OzSTAR is funded by Swinburne University of Technology.

DATA AVAILABILITY

Information to access to the raw simulation data underlying this article can be found at <https://bradgreig.github.io/large-volume-data/>. All other data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018a, *Nature*, 555, 67
 Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018b, *Nature*, 564, E35
 Bradley R. F., Tauscher K., Rapetti D., Burns J. O., 2019, *ApJ*, 874, 153
 Chapman E., Jelić V., 2019, preprint ([arXiv:1909.12369](https://arxiv.org/abs/1909.12369))
 Cohen A., Fialkov A., Barkana R., 2018, *MNRAS*, 478, 2193
 Cooray A., 2005, *MNRAS*, 363, 1049
 Datta A., Bowman J. D., Carilli C. L., 2010, *ApJ*, 724, 526
 Datta K. K., Mellema G., Mao Y., Iliev I. T., Shapiro P. R., Ahn K., 2012, *MNRAS*, 424, 1877
 Datta K. K., Jensen H., Majumdar S., Mellema G., Iliev I. T., Mao Y., Shapiro P. R., Ahn K., 2014, *MNRAS*, 442, 1491
 DeBoer D. R. et al., 2017, *PASP*, 129, 045001
 Draine B. T., Miralda-Escudé J., 2018, *ApJ*, 858, L10
 Eastwood M. W. et al., 2019, *AJ*, 158, 84

Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004a, *ApJ*, 613, 1
 Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004b, *ApJ*, 613, 16
 Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181
 Ghara R., Datta K. K., Choudhury T. R., 2015, *MNRAS*, 453, 3143
 Ghara R. et al., 2020, *MNRAS*, 493, 4728
 Ghara R., Giri S. K., Ciardi B., Mellema G., Zaroubi S., 2021, *MNRAS*, 503, 4551
 Giri S. K., D'Aloisio A., Mellema G., Komatsu E., Ghara R., Majumdar S., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 058
 Gnedin N. Y., Ostriker J. P., 1997, *ApJ*, 486, 581
 Gnedin N. Y., Shaver P. A., 2004, *ApJ*, 608, 611
 Greig B., Mesinger A., 2015, *MNRAS*, 449, 4246
 Greig B., Bolton J. S., Wyithe J. S. B., 2011, *MNRAS*, 418, 1980
 Greig B., Mesinger A., Koopmans L. V. E., 2020, *MNRAS*, 491, 1398
 Greig B., Trott C. M., Barry N., Mutch S. J., Pindor B., Webster R. L., Wyithe J. S. B., 2021a, *MNRAS*, 500, 5322
 Greig B. et al., 2021b, *MNRAS*, 501, 1
 Gupta Y. et al., 2017, *Curr. Sci.*, 113, 707
 Hills R., Kulkarni G., Meerburg P. D., Puchwein E., 2018, *Nature*, 564, E32
 Hutter A., Watkinson C. A., Seiler J., Dayal P., Sinha M., Croton D. J., 2020, *MNRAS*, 492, 653
 Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, *MNRAS*, 439, 725
 Jensen H., Majumdar S., Mellema G., Lidz A., Iliev I. T., Dixon K. L., 2016, *MNRAS*, 456, 66
 Kaur H. D., Gillet N., Mesinger A., 2020, *MNRAS*, 495, 2354
 Kim H.-S., Wyithe J. S. B., Park J., Poole G. B., Lacey C. G., Baugh C. M., 2016, *MNRAS*, 455, 4498
 Koopmans L. et al., 2015, *Proc. Sci., Advancing Astrophysics with the Square Kilometre Array (AASKA14)*. SISSA, Trieste, PoS#215
 Lacey C., Cole S., 1993, *MNRAS*, 262, 627
 La Plante P., Battaglia N., Natarajan A., Peterson J. B., Trac H., Cen R., Loeb A., 2014, *ApJ*, 789, 31
 Line J., 2022, *J. Open Source Softw.*, 7, 3676
 Liu A., Parsons A. R., Trott C. M., 2014a, *Phys. Rev. D*, 90, 023018
 Liu A., Parsons A. R., Trott C. M., 2014b, *Phys. Rev. D*, 90, 023019
 Madau P., Meiksin A., Rees M. J., 1997, *ApJ*, 475, 429
 Majumdar S., Pritchard J. R., Mondal R., Watkinson C. A., Bharadwaj S., Mellema G., 2018, *MNRAS*, 476, 4007
 Mellema G. et al., 2013, *Exp. Astron.*, 36, 235
 Mertens F. G. et al., 2020, *MNRAS*, 493, 1662
 Mesinger A., Furlanetto S., 2007, *ApJ*, 669, 663
 Mesinger A., Furlanetto S., Cen R., 2011, *MNRAS*, 411, 955

- Mondal R., Bharadwaj S., Majumdar S., Bera A., Acharyya A., 2015, *MNRAS*, 449, L41
- Mondal R., Bharadwaj S., Majumdar S., 2016, *MNRAS*, 456, 1936
- Mondal R., Bharadwaj S., Majumdar S., 2017, *MNRAS*, 464, 2992
- Mondal R., Bharadwaj S., Datta K. K., 2018, *MNRAS*, 474, 1390
- Mondal R. et al., 2020, *MNRAS*, 498, 4178
- Morales M. F., Hewitt J., 2004, *ApJ*, 615, 7
- Morales M. F., Wyithe J. S. B., 2010, *ARA&A*, 48, 127
- Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, *ApJ*, 752, 137
- Murray S. G., Trott C. M., 2018, *ApJ*, 869, 25
- Nasirudin A., Murray S. G., Trott C. M., Greig B., Joseph R. C., Power C., 2020, *ApJ*, 893, 118
- Park J., Mesinger A., Greig B., Gillet N., 2019, *MNRAS*, 484, 933
- Parsons A. R. et al., 2010, *AJ*, 139, 1468
- Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, *ApJ*, 756, 165
- Planck Collaboration VI, 2020, *A&A*, 641, A6
- Pober J. C. et al., 2016, *ApJ*, 819, 8
- Pritchard J. R., Loeb A., 2012, *Rep. Prog. Phys.*, 75, 086901
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Scoccimarro R., 1998, *MNRAS*, 299, 1097
- Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, *A&A*, 345, 380
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Singh S., Subrahmanyan R., 2019, *ApJ*, 880, 26
- Singh S. et al., 2022, *Nat. Astron.*, 6, 607
- Sobacchi E., Mesinger A., 2014, *MNRAS*, 440, 1662
- The HERA Collaboration, 2022a, *ApJ*, 924, 51
- The HERA Collaboration, 2022b, *ApJ*, 925, 221
- Thyagarajan N. et al., 2013, *ApJ*, 776, 6
- Thyagarajan N. et al., 2015a, *ApJ*, 804, 14
- Thyagarajan N. et al., 2015b, *ApJ*, 807, L28
- Tingay S. J. et al., 2013, *Publ. Astron. Soc. Aust.*, 30, 7
- Tozzi P., Madau P., Meiksin A., Rees M. J., 2000, *ApJ*, 528, 597
- Trott C. M., Wayth R. B., Tingay S. J., 2012, *ApJ*, 757, 101
- Trott C. M. et al., 2020, *MNRAS*, 493, 4711
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- Vedantham H., Shankar N. U., Subrahmanyan R., 2012, *ApJ*, 745, 176
- Wayth R. et al., 2018, *Publ. Astron. Soc. Aust.*, 35, 33
- Zahn O., Mesinger A., McQuinn M., Trac H., Cen R., Hernquist L. E., 2011, *MNRAS*, 414, 727
- Zarka P., Girard J. N., Tagger M., Denis L., 2012, in Boissier S., de Laverny P., Nardetto N., Samadi R., Valls-Gabaud D., Wozniak H., eds, SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, French Society of Astronomy and Astrophysics, Paris, p. 687

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.