

# Dark-ages reionization and galaxy formation simulation – XII. Bubbles at dawn

Paul M. Geil,<sup>1,2★</sup> Simon J. Mutch,<sup>1</sup> Gregory B. Poole,<sup>1</sup> Alan R. Duffy,<sup>3</sup>  
Andrei Mesinger<sup>4</sup> and J. Stuart B. Wyithe<sup>1,2</sup>

<sup>1</sup>*School of Physics, The University of Melbourne, Parkville, Victoria 3010, Australia*

<sup>2</sup>*ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)*

<sup>3</sup>*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*

<sup>4</sup>*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

Accepted 2017 July 18. Received 2017 June 23; in original form 2017 April 17

## ABSTRACT

The direct detection of regions of ionized hydrogen (H II) has been suggested as a promising probe of cosmic reionization. Observing the redshifted 21-cm signal of hydrogen from the epoch of reionization (EoR) is a key scientific driver behind new-generation, low-frequency radio interferometers. We investigate the feasibility of combining low-frequency observations with the Square Kilometre Array and near infra-red survey data of the *Wide-Field Infrared Survey Telescope* to detect cosmic reionization by imaging H II bubbles surrounding massive galaxies during the cosmic dawn. While individual bubbles will be too small to be detected, we find that by stacking redshifted 21-cm spectra centred on known galaxies, it will be possible to directly detect the EoR at  $z \sim 9\text{--}12$ , and to place qualitative constraints on the evolution of the spin temperature of the intergalactic medium (IGM) at  $z \gtrsim 9$ . In particular, given a detection of ionized bubbles using this technique, it is possible to determine if the IGM surrounding them is typically in absorption or emission. Determining the globally averaged neutral fraction of the IGM using this method will prove more difficult due to degeneracy with the average size of H II regions.

**Key words:** galaxies: high-redshift – intergalactic medium – dark ages, reionization, first stars.

## 1 INTRODUCTION

Cosmic hydrogen is believed to have been reionized by ultraviolet (UV) radiation produced by stars and quasars. The period from the formation of the first ionizing sources to when the intergalactic medium (IGM) was completely ionized is commonly known as the epoch of reionization (EoR). However, due to formidable challenges in observation and simulation, our knowledge of this process is lacking. Knowing how reionization occurred, in both time and space, would dramatically improve our understanding not only of the evolution and properties the IGM, but also of the formation and role of the ionizing sources responsible during this period.

Observations of high-redshift sources and the cosmic microwave background (CMB) have allowed some constraints to be placed on the timing and duration of the EoR. For example, Gunn-Peterson absorption troughs in quasar Lyman  $\alpha$  spectra set a lower limit for the end of reionization at  $z \sim 6$  (Fan, Carilli & Keating 2006; Mortlock et al. 2011). Additionally, CMB observations provide a measure of the total optical depth to electron scattering. Since this is an integrated quantity from the surface of last scattering ( $z \sim 1100$ ),

it cannot, on its own, distinguish between different reionization histories. However, depending on the model of reionization adopted, the average redshift at which reionization is half complete is found to lie between  $z = 7.8$  and  $8.8$  (Planck Collaboration XLVII 2016b). Recent analysis by Greig et al. (2017) implies that reionization is not yet complete by  $z = 7.1$ , with the volume-weighted IGM neutral fraction constrained to  $0.40^{+0.41}_{-0.32}$  at  $2\sigma$ .

A far more promising observational strategy to constrain reionization is to directly measure the emission from the 21-cm spin-flip transition of neutral hydrogen (H I). Due to the cosmic expansion, the frequency of this radiation is now  $<200$  MHz. Various experiments are underway or planned to measure the cosmic 21-cm signal as a function of frequency (and therefore redshift, time or distance) utilizing different methods. One approach is to measure the spatially averaged global signal using a single-dipole antenna, e.g. EDGES (Bowman, Rogers & Hewitt 2008), DARE (Burns et al. 2012) and SARAS (Patra et al. 2013). Another is to measure the signal's spatial fluctuations interferometrically (using e.g. LOFAR,<sup>1</sup> GMRT,<sup>2</sup>

\* E-mail: [geil.p@unimelb.edu.au](mailto:geil.p@unimelb.edu.au)

<sup>1</sup> <http://www.lofar.org>

<sup>2</sup> <http://www.ncra.tifr.res.in/ncra/gmrt>

PAPER,<sup>3</sup> MWA,<sup>4</sup> HERA,<sup>5</sup> SKA<sup>6</sup>). For some instruments the latter approach can yield both high-resolution tomographic images of the ionized structure and statistical measurements (such as the 21-cm power spectrum) allowing us to learn about the properties of the reionization sources and sinks in far greater detail than simple timing estimates. For reviews of the EoR science and 21-cm detection experiments, see e.g. Morales & Wyithe (2010) and Koopmans et al. (2015).

In this work we use simulations to investigate structures of ionized hydrogen (H II) surrounding the first galaxies during the early stages of the EoR ( $z \gtrsim 9$ ). During this era – known also as the cosmic dawn – these regions appear as isolated ‘bubbles’. We begin by discussing the ionized regions associated with simulation analogues of the highest-known redshift galaxy to date (GN-z11). We then move on to consider the wider population of bubbles in our simulation, establishing a relationship between their size, and redshift and luminosity of the brightest galaxy within them. We apply this simulation-based empirical relationship to explore the utility of an image regime-based EoR detection strategy that synergizes the proposed *Wide-Field Infrared Survey Telescope’s (WFIRST)* High Latitude Survey (HLS) and deep integrations of the redshifted 21-cm signal using the planned low-frequency Square Kilometre Array (SKA1-LOW). Our direct detection strategy is similar to those proposed targeting regions of ionized hydrogen surrounding high-luminosity quasars (see e.g. Kohler et al. 2005; Wyithe, Loeb & Barnes 2005; Geil & Wyithe 2008) but is able to push the detection redshift beyond what is possible using quasars alone due to their relatively low population at  $z > 8$ . Other techniques for probing individual sources have been presented, both midway through the EoR and at very high redshift ( $z \sim 15$ ), such as visibility-based methods using matched filtering (e.g. Datta, Bharadwaj & Choudhury 2007; Datta et al. 2012; Majumdar, Bharadwaj & Choudhury 2012; Ghara, Choudhury & Datta 2016), and using imaging (Ghara et al. 2017). Some of these works also assess the prospects of constraining properties of the high-redshift IGM (such as its globally averaged neutral fraction) and the sources responsible for its reionization.

This paper is structured as follows. In Section 2 we give a brief overview of the DRAGONS simulation used in this paper. Section 3 explains the motivation behind this work and presents our bubble size–galaxy redshift and luminosity relation. Section 4 describes our detection strategy and presents our detectability results. Section 5 explores simple methods to constrain the spin temperature and globally averaged ionization state of the high-redshift IGM. We address some additional details that may potentially impact our results in Section 6 before presenting a summary in Section 7. We include an Appendix containing supporting material detailing the model fitting, UV luminosity functions (UVLFs) and instrumental noise estimates used in this work. All globally averaged quantities (e.g. neutral fraction) are volume weighted, and distances are given in comoving units unless stated otherwise. Absolute magnitudes used throughout are given in the AB system (Oke & Gunn 1983), are intrinsic and have been calculated using the methodology described in Liu et al. (2016), assuming a standard Salpeter (1955) initial mass function with upper and lower mass limits of  $0.1 M_{\odot}$  and  $120 M_{\odot}$ , respectively. Our choice of cosmology is the stan-

dard spatially flat *Planck*  $\Lambda$ CDM cosmology (Planck Collaboration XIII 2016b) ( $h, \Omega_m, \Omega_b, \Omega_{\Lambda}, \sigma_8, n_s$ ) = (0.678, 0.308, 0.0484, 0.692, 0.815, 0.968).

## 2 THE DRAGONS SIMULATION

The Dark-ages, Reionization And Galaxy-formation Observables from Numerical Simulations (DRAGONS<sup>8</sup>) project was specifically designed to study the formation of the first galaxies and cosmic reionization. It integrates a semi-numerical calculation of reionization (21CMFAST) within a semi-analytic model of galaxy formation (MERAXES) built upon an  $N$ -body simulation (*Tiamat*). This gives a self-consistently coupled reionization model which accounts for feedback due to both supernovae (SN) and the ionizing UV background from stars within galaxies. A unique feature of DRAGONS is that it utilizes *horizontal* rather than vertical dark matter halo merger trees. This allows it to correctly simulate how galaxies influence each others’ evolution by way of their ionizing flux. *Tiamat* has a sufficiently large volume (cube of sides 100 Mpc in length) to investigate cosmic evolution while still achieving a mass resolution approaching the atomic cooling mass threshold. *Tiamat* also has excellent temporal resolution with 100 equally spaced snapshots between redshifts 5 and 35, giving a cadence of about 11 Myr. This means the stochastic effects of star formation and SN feedback on reionization are accurately captured. Complete descriptions of *Tiamat* and MERAXES are given in Poole et al. (2016) (Paper I) and Mutch et al. (2016a) (Paper III), respectively, while details of the 21CMFAST algorithm are described in Mesinger, Furlanetto & Cen (2011). Geil et al. (2016) (hereafter Paper V) investigate the effect of galaxy-formation physics on the morphology and statistical signatures of reionization.

The MERAXES model used in this work is the fiducial model described in Papers III and V. This model has been calibrated so as to reproduce the observed evolution of the galaxy stellar mass function from  $z = 5$  to 7 (see Paper III) and the latest *Planck* optical depth measurements (Planck Collaboration XIII 2016a). All output fields (e.g. density, stellar mass and ionization fraction) have been regularly gridded over  $512^3$  voxels.

## 3 H II REGIONS SURROUNDING THE FIRST GALAXIES

### 3.1 GN-z11 analogues: DR-1 and DR-2

Motivated by the identification of the surprisingly bright and massive galaxy GN-z11 at  $z = 11.1$  by Oesch et al. (2016), Mutch et al. (2016b) (Paper VI) investigate the origin and fate of such objects using DRAGONS. Two analogue galaxies of similar luminosity and stellar mass to GN-z11, labelled DR-1 and DR-2, were found within the *Tiamat* volume and show excellent agreement with all available observationally derived properties of this object. Maintaining this motivation, here we briefly summarize aspects of these objects’ impact on the IGM in terms of reionization.

With the *Tiamat* volume gridded to  $512^3$ , the voxels containing DR-1 and DR-2 are first fully ionized at  $z_{\text{ion}} = 17.8$  and 17.1, respectively. While these objects were not the first sources to begin ionizing the IGM in our simulation (on scales corresponding to this grid resolution), they were among the first, with the majority of voxels being ionized after  $z \approx 7.9$  (approximately corresponding to

<sup>3</sup> <http://eor.berkeley.edu>

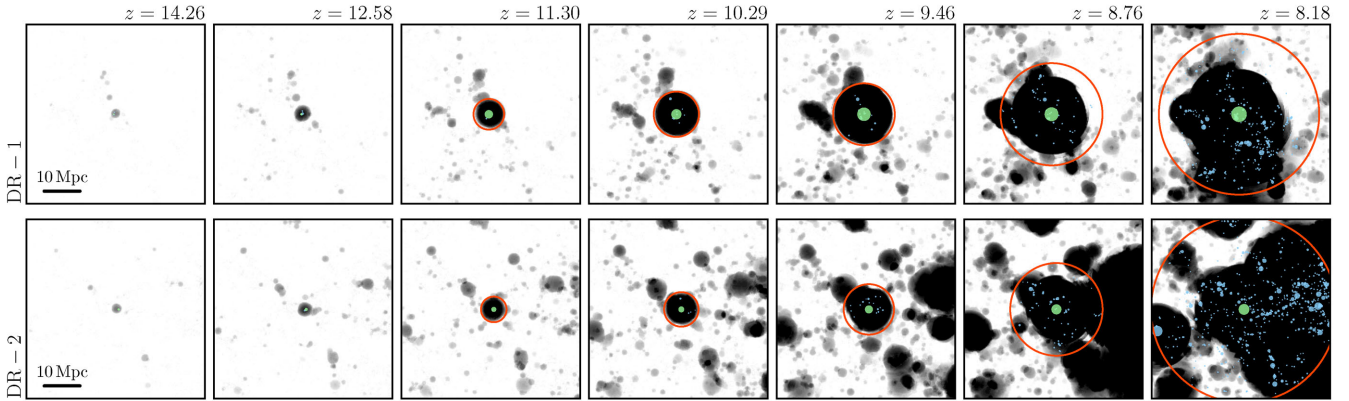
<sup>4</sup> <http://www.mwatelescope.org>

<sup>5</sup> <http://reionization.org>

<sup>6</sup> <http://www.skatelescope.org>

<sup>7</sup> <http://wfIRST.gsfc.nasa.gov>

<sup>8</sup> <http://dragons.ph.unimelb.edu.au>



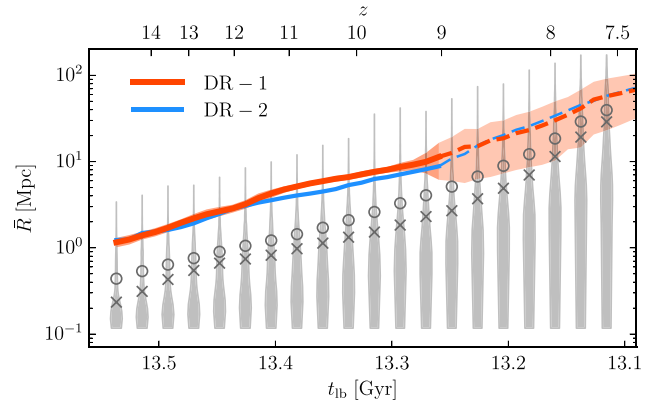
**Figure 1.** Slices through the ionization fields surrounding DR-1 (top) and DR-2 (bottom) at selected redshifts showing the evolution of bubble size. The positions of DR-1 and DR-2 are indicated by the central (green) filled circles. The projected positions of other galaxies within the average radius of the bubble (shown by the red circle for panels with  $z < 12$ ) are indicated by the other (blue) filled circles. The area of each galaxy’s marker is proportional to its UV luminosity. Each slice is 5 Mpc deep.

the redshift at which half the hydrogen in the IGM is ionized). This is not surprising given these objects’ early star formation histories (see Paper VI) and the fact that they are found in highly overdense regions.

The ionized bubbles surrounding DR-1 and DR-2 (and the other less massive and less luminous galaxies within them) at  $z = 11.1$  are approximately spherical. This is due to the lack of overlapping bubbles surrounding other galaxies in their vicinity. At this redshift both DR-1 and DR-2 lie close to the centroid of their bubbles. We estimate individual bubble size using a three-dimensional ray tracing technique, centred upon the brightest galaxy in the bubble, which measures the distance to an ionization phase transition (demarcated by a step to a voxel that is more than 50 per cent neutral) in  $\geq 10^3$  randomly chosen directions. In the case of DR-1 and DR-2 at  $z = 11.1$  this provides an accurate and precise estimate of bubble radius. However, at later times (when there is overlap and the brightest galaxy in the bubble may be off-centre) the resulting sampled radius distribution has higher variance (which can be used to mark the approximate transition from an isolated bubble to an overlapping region). Using this method we find that the average diameters of the bubbles surrounding DR-1 and DR-2 at  $z = 11.1$  are  $\approx 10$  and 8 Mpc, respectively. At this redshift the globally averaged neutral fraction,  $\bar{x}_{\text{HI}}$ , of our fiducially modelled IGM is 0.976, hence these two bubbles alone (out of  $\gtrsim 600$ ) make up just under 3 per cent of the total ionized volume.

Fig. 1 shows zoomed-in slices through the ionization fields surrounding DR-1 and DR-2 at selected redshifts, showing the evolution of bubble size. The positions of DR-1 and DR-2 are indicated by the central (green) filled circles. The projected positions of other galaxies (as faint as  $M_{\text{UV}} = -17.25$ ) within the average radius of the bubble (shown by the red circle for panels with  $z < 12$ ) are indicated by the other (blue) filled circles. The area of each galaxy’s marker is proportional to its UV luminosity. By visual inspection these bubbles cease to be isolated regions, and are also driven by many less luminous galaxies, from  $z \approx 9$ .

In order to compare the bubbles surrounding DR-1 and DR-2 to others in the simulation, Fig. 2 shows where they lie in the size distribution of all ionized regions in the simulation as a function of redshift (and lookback time). The average size of the bubbles surrounding DR-1 and DR-2 is shown by the thick red and thinner blue lines, respectively. The shaded region indicates the  $\pm 1\sigma$  range in radius for DR-1 (the uncertainty for DR-2 is not shown, but is

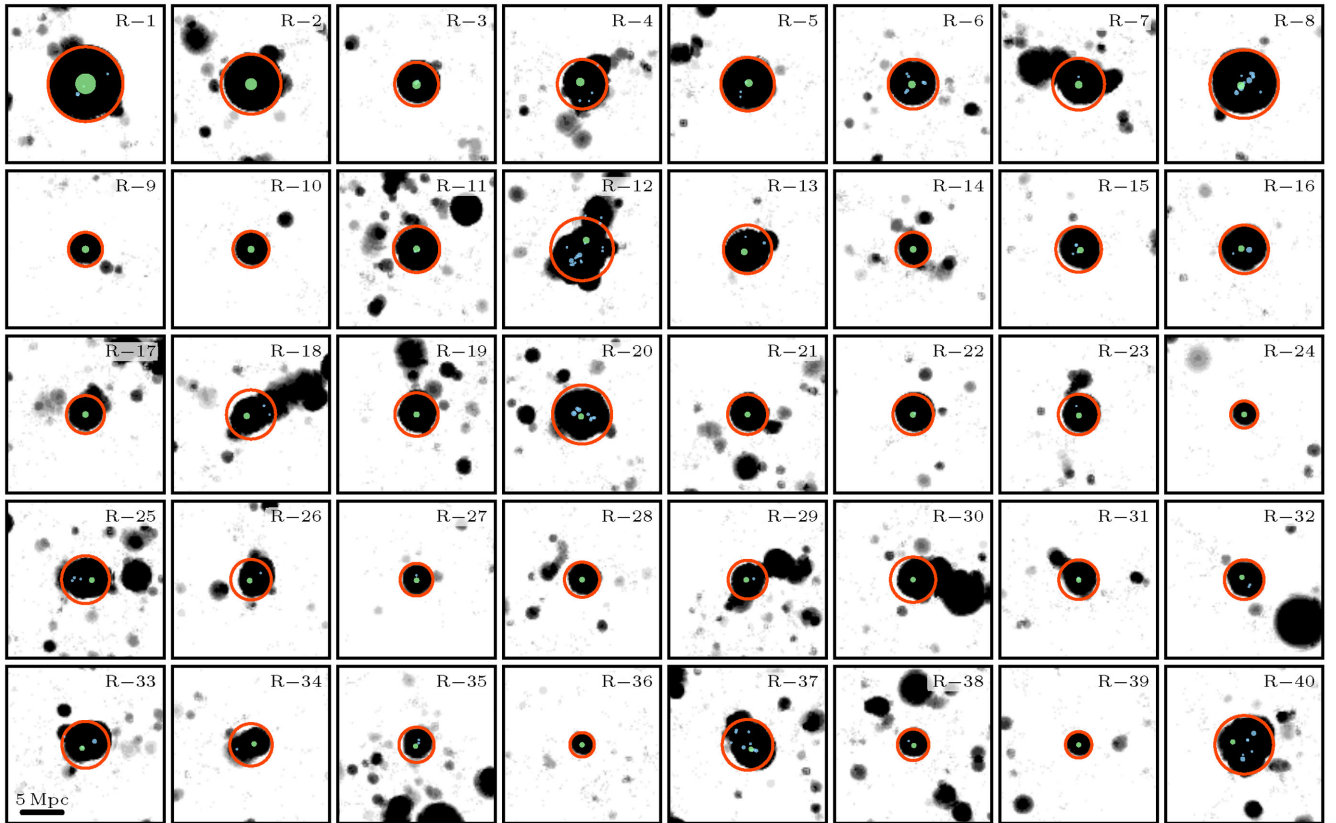


**Figure 2.** Evolution of the size of bubbles surrounding DR-1 (red) and DR-2 (blue) as a function of redshift and lookback time. The red shaded region indicates the  $\pm 1\sigma$  range in radius for DR-1, calculated using the ray tracing technique described in Section 3.1. Dashed lines are used when the error in radius is more than half the radius of the bubble (marking the approximate transition from an isolated bubble to an overlapping region). For comparison, the probability densities of the size of other ionized regions in the simulation at selected redshifts are shown by the violin plot (grey; with circles and crosses indicating their means and medians, respectively).

similar to that of DR-1), calculated using the ray tracing technique described above. Dashed extensions of the lines show when the error in radius is more than half the radius of the bubble, marking the approximate transition from an isolated bubble to an overlapping region. For comparison, the sizes of other ionized regions in the simulation<sup>9</sup> are shown by the distributions (with circles and crosses indicating their means and medians, respectively).

<sup>9</sup> We calculate the bubble size distribution for each snapshot using the Monte Carlo method described in Mesinger & Furlanetto (2007). This is the same as the ray tracing technique used to estimate individual bubble size described earlier; however, in this method, a voxel that is more than 50 per cent ionized is first randomly selected from the full gridded simulation volume and its distance from an ionization phase transition in a randomly chosen direction is recorded. This is repeated  $10^7$  times to form a probability distribution function of region size. This methodology also provides an approximate measure of the mean-free path of ionizing photons inside ionized regions.





**Figure 3.** Slices through the ionization fields of the first 40 (unique) bubbles surrounding the brightest galaxies at  $z = 11.1$ . The projected positions of galaxies within the average radius of the bubble (shown by the red circle) are indicated by the filled circles (the brightest in green). Each slice is 1.5 Mpc deep and has been centred on the centre of luminosity of the galaxies the bubble contains.

### 3.2 Bubble size–luminosity relation

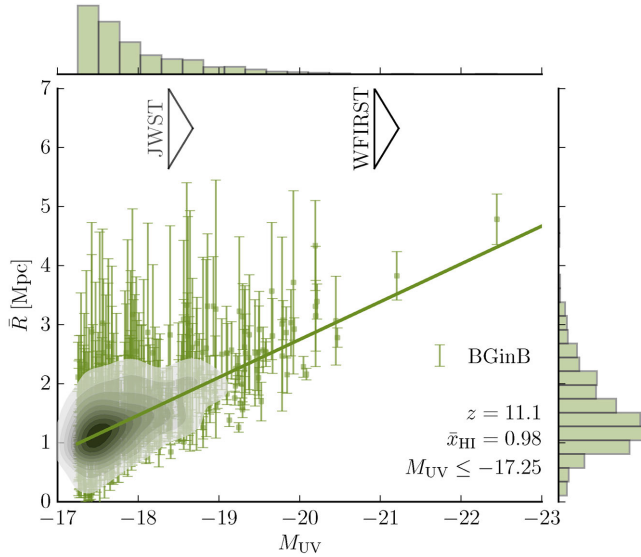
In this section we turn our attention to the population of bubbles surrounding a deep selection of galaxies in our simulation. Our main objective here is to investigate the expected connection between the size of such regions (in terms of mean radius,  $\bar{R}$ ) and the luminosity of the brightest galaxy within them (in terms of their intrinsic absolute UV magnitude,  $M_{UV}$ ), aiming to establish a simple relationship between these properties as a function of redshift. We do so anticipating its use in Section 4, where we examine the prospects of detecting ionized regions at high redshift. Note that our galaxy number density predictions are based on intrinsic luminosities and do not include dust attenuation as this is not expected to be significant at such high redshifts ( $z \geq 9$ ). This also maintains the good agreement between the BLUE TIDES UVLFs used in this work (Waters et al. 2016) and the results of Oesch et al. (2016).

Before demonstrating our  $\bar{R}-M_{UV}$  fitting procedure we show a sample of zoomed-in slices through the ionization fields of the first 40 (unique) bubbles surrounding the brightest galaxies at  $z = 11.1$  in Fig. 3. Note that for the purpose of detecting the EoR by stacking bubble 21-cm spectra, we are interested in the relationship between bubble size and the luminosity of the *brightest galaxy in the bubble*. Hence, only one datum contributes to the  $\bar{R}-M_{UV}$  model fitting for each bubble and therefore the bubbles shown in Fig. 3 are unique. In this case, since DR-1 and DR-2 are at a significant distance from one another, they happen to be the brightest galaxies in R-1 and R-2, respectively. In general, however, there is no one-to-one correspondence between bubbles and galaxies due to clustering. The average radius of each bubble has been estimated using the individ-

ual bubble method described in Section 3.1. The projected positions of galaxies within the average radius of the bubble (shown by the red circle) are indicated by the filled circles (the brightest in green). Each slice is 1.5 Mpc deep and, for aesthetics, has been centred on the centre of luminosity of the galaxies the bubble contains. We include this figure in order to demonstrate the variation in geometry of these regions at this redshift.

The plot of average bubble radius against the absolute UV magnitude of the brightest galaxy within each bubble (considering galaxies brighter than  $M_{UV} = -17.25$  only) at  $z = 11.1$  is shown in Fig. 4. The error bars represent the  $\pm 1\sigma$  range in radius. Histograms on the top and right axes indicate the marginalized distributions of UV magnitude and bubble radius, respectively. For reference, the  $5\sigma$  detection limit of the *WFIRST* HLS ( $m_{UV} = 26.75$ ; Spergel et al. 2013) and the  $8\sigma$  detection limit of a wide-field survey using the *James Webb Space Telescope* (*JWST*),<sup>10</sup>  $m_{UV} = 29.3$ ; Mason, Trenti & Treu 2015) at this redshift are indicated. We use this result, and similar results at other redshifts (again, considering galaxies brighter than  $M_{UV} = -17.25$  only), to perform an error-weighted Markov Chain Monte Carlo (MCMC) parameter estimation to the linear  $\bar{R}-M_{UV}$  model,  $\bar{R} = a_0 + a_1 M_{UV}$ , at 18 redshifts between  $z \sim 9$  and 12. We also fit for an estimate of the variance in  $a_0$ ,  $\sigma_0^2$ . Specific results for  $z = 9.2, 10.2$  and  $11.1$  ( $\bar{x}_{H\text{I}} \approx 0.85, 0.94$  and  $0.98$ , respectively) are given in Table 1. While one may expect a *non-linear*  $\bar{R}-M_{UV}$  relationship (e.g.  $R \propto L_{UV}^{1/3} \propto 10^{-0.4M_{UV}/3}$  for a

<sup>10</sup> <http://www.jwst.nasa.gov/>



**Figure 4.** Bubble radius versus absolute UV magnitude of the brightest galaxy it contains (labelled ‘BGinB’ for brightest galaxy in bubble) at  $z = 11.1$  for  $M_{UV} \leq -17.25$ . The error bars show the  $\pm 1\sigma$  range in radius, calculated using the ray tracing technique described in Section 3.1. The solid line is the best error-weighted linear fit. The large arrows indicate the *WFIRST*  $5\sigma$  and *JWST*  $8\sigma$  detection limits at this redshift (see text for details). Histograms on the top and right axes indicate the marginalized distributions of UV magnitude and bubble radius, respectively.

**Table 1.** Resulting best-fitting error-weighted  $\bar{R}$ - $M_{UV}$ -model parameter values for selected redshifts/neutral fractions. Confidence limits are determined by the 68 per cent confidence intervals of the marginalized distributions.

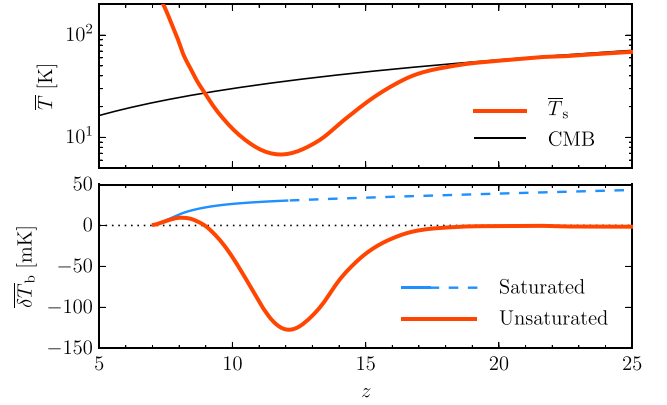
$z$	$\bar{x}_{HI}$	$a_1$ [Mpc mag $^{-1}$ ]	$a_0$ [Mpc]	$\sigma_0^2$ [Mpc $^2$ ]
9.2	0.85	$-1.28^{+0.03}_{-0.03}$	$-20.66^{+0.63}_{-0.63}$	$0.84^{+0.05}_{-0.05}$
10.1	0.94	$-0.82^{+0.02}_{-0.02}$	$-13.00^{+0.45}_{-0.44}$	$0.27^{+0.02}_{-0.02}$
11.1	0.98	$-0.64^{+0.02}_{-0.02}$	$-10.06^{+0.41}_{-0.42}$	$0.11^{+0.01}_{-0.01}$

cosmological Strömgren sphere generated by a source of UV luminosity,  $L_{UV}$ ), many other galaxies in the neighbourhood of the brightest galaxy in each bubble enhance the local ionizing photon budget. Bias and clustering of these sources conspire to complicate this relationship. We choose to fit a linear model for both simplicity and the fact that it describes the luminosity enhancement well (see Appendix A).

Using these results we perform non-linear fits in redshift for  $a_1$ ,  $a_0$  and  $\sigma_0^2$  (the bespoke functional form and best-fitting parameters are given in Appendix A) to enable us to calculate estimates for  $\bar{R}$  and  $\sigma_{\bar{R}}$  as a function of magnitude and redshift. These provide the statistical description for the mock bubble size distributions (which we assume to be Gaussian) we utilize in Section 4.3.

#### 4 DETECTABILITY

Having established a functional relationship between the typical size of bubbles as a function of both the luminosity of the brightest galaxy within them and its redshift, we now investigate prospects for their detectability – and therefore direct evidence of cosmic reionization. The two principal competing components we consider to be at play



**Figure 5.** Top panel: evolution of the volume-averaged spin temperature according to the Evolution Of 21 cm Structure simulation (Mesinger et al. 2016). The evolution of the CMB temperature is shown for comparison. Bottom panel: corresponding volume-averaged 21-cm differential brightness temperature for both the saturated and unsaturated signal case. The dashed extension of the saturated case beyond  $z \approx 12$  indicates where the IGM is unlikely to have been fully heated.

here are the strength of the cosmic 21-cm signal and instrumental noise. We describe the formulation of both in Sections 4.1 and 4.2, respectively, before presenting our detection strategy in Section 4.3.

#### 4.1 Cosmic 21-cm signal

The relevant cosmic signal is the spatially dependent 21-cm differential brightness temperature,  $\delta T_b$ , between hydrogen gas and the CMB along the line of sight (for a detailed discussion of the fundamental physics of the 21-cm line, see e.g. Furlanetto, Oh & Briggs 2006). For  $z \gg 1$ ,  $\delta T_b$  can be written as

$$\delta T_b \approx 27x_{HI}(1 + \delta) \left(1 - \frac{T_\gamma}{T_s}\right) \left(\frac{1+z}{10}\right)^{1/2} \left(\frac{0.15}{\Omega_m h^2}\right)^{1/2} \times \left(\frac{\Omega_b h^2}{0.023}\right) \text{ mK}, \quad (1)$$

where  $\delta = \delta(\mathbf{x}, z) \equiv \rho(\mathbf{x}, z)/\bar{\rho}(z) - 1$  is the local dark matter overdensity at position  $\mathbf{x}$  and redshift  $z$ ,  $x_{HI} = x_{HI}(\mathbf{x}, z)$  the local neutral fraction,  $T_s = T_s(\mathbf{x}, z)$  the local spin temperature, and  $T_\gamma = 2.73(1+z)$  K the CMB temperature. By using this formulation we ignore redshift-space distortions. When  $T_s = T_\gamma$  the 21-cm signal from the IGM vanishes. Similarly, as reionization progress ( $x_{HI} \rightarrow 0$ ), the 21-cm signal diminishes. When  $T_s < T_\gamma$  ( $T_s > T_\gamma$ ) the 21-cm signal appears in absorption (emission). In the post-heating regime, where X-rays heat the IGM and the Lyman  $\alpha$  background acts to decouple the 21-cm transition from the CMB (such that  $T_s \gg T_\gamma$ ), the 21-cm signal *saturates* (since  $1 - T_\gamma/T_s \rightarrow 1$  in equation 1) and it appears in emission. This is demonstrated in the top panel of Fig. 5, which shows the evolution of the volume-averaged spin temperature,  $\bar{T}_s$ , according to the Evolution Of 21 cm Structure (EOS) simulation by Mesinger, Greig & Sobacchi (2016) using their ‘bright galaxies’ model. In this model, reionization is dominated by galaxies inside  $> 10^{10} M_\odot$  haloes (roughly corresponding to  $M_{UV} \lesssim -17$ ). The bottom panel shows the corresponding volume-averaged 21-cm differential brightness temperature for both the unsaturated and saturated signal case (the dashed extension of the saturated case beyond  $z \approx 12$  indicates where the IGM is unlikely to have been fully heated.). The EOS simulation

incorporates extremely efficient SNe feedback and closely matches the global reionization history of our fiducial model. For computational efficiency, we apply the temperature differential factor in equation (1) homogeneously to our  $\delta T_b$  fields (which have been calculated assuming signal saturation) using the EOS spin temperature model (i.e. we simply adjust the mean temperature).

#### 4.2 Instrumental noise

We simulate the instrumental noise of SKA1-LOW based on the imaging performance results provided by SKAO Science Team (2015) (hereafter SKA2015). As our detection strategy involves using line-of-sight redshifted 21-cm spectra, the spatial structure of the noise (i.e. in the sky plane) does not concern us. Rather, all that is required is the rms noise,  $\sigma_N$ , for an observation made at frequency  $\nu$ , with frequency channel width  $\Delta\nu$ , integrated over time  $t_{\text{int}}$ , and smoothed using a synthesized beam of full width at half-maximum (FWHM)  $\theta_b$ . More specific details are included in Appendix B.

#### 4.3 Stack-averaged line-of-sight 21-cm spectra

##### 4.3.1 Individual verses stack-averaged bubbles

In order to provide a sense of the relative brightness temperatures of signals and noise involved, consider the bubble surrounding DR-1 at  $z = 11.1$  (117.6 MHz). According to our simulation the globally averaged neutral fraction of the surrounding IGM is 0.98. Therefore, using equation (1) and assuming saturation, the bubble appears as a near-spherical,  $\approx 30$  mK deep ‘hole’ of radius  $\approx 5$  Mpc. The instrumental noise, smoothed in both the sky plane and frequency space on a scale equal to the radius of the bubble ( $\approx 2$  arcmin and  $\approx 240$  kHz, respectively; large enough to reduce the instrumental noise without completely smoothing out the bubble<sup>11</sup>), has an rms of  $\sigma_N \approx 100$  mK. Not assuming saturation gives an  $\approx 100$  mK signal. The situation improves at  $z = 9.2$ , but detection is still unlikely, with an  $\approx 23$  mK saturated signal ( $\approx 5$  mK unsaturated) hidden in noise with  $\sigma_N \approx 10$  mK.

Given the unlikely prospect of detecting *individual* bubbles surrounding galaxies, we turn our attention to stack-averaging multiple line-of-sight 21-cm spectra in order to improve the signal-to-noise ratio (SNR). As an example, we look at the case using the deep and wide near-infrared HLS by *WFIRST* to obtain sky position, redshift and UV magnitude data of galaxies that lie within the SKA deep integration field.<sup>12</sup> As *WFIRST*’s planned spectroscopic survey will not be sufficiently deep to detect the same, relatively faint, galaxy candidates as identified by the HLS imaging survey, accurate redshift determination will require follow-up grism spectroscopy. Redshifts would be estimated by fitting a spectral energy distribution to the spectra, as was done by Oesch et al. (2016) for GN-z11’s spectra from the Wide Field Camera 3 (WFC3) slitless grism on the *Hubble Space Telescope* (*HST*). Furthermore, without emission

lines, these galaxies will need to be Lyman break galaxies (as this break is by far the strongest feature in the spectra). The presence of emission lines would likely improve redshift determination considering the high spectral resolution of *WFIRST*’s grism ( $R = 600$ ; Spergel et al. 2013). At the redshifts investigated in this work, it is reasonable to expect that most of the detectable galaxies will exhibit a Lyman break due to the low ionization fraction of the intervening IGM. We also note that spectroscopic follow-up could also be performed using other instruments, e.g. *JWST*. These sky position, redshift and UV magnitude data are then used to stack-average the line-of-sight redshifted 21-cm spectra centred on each target galaxy. Ideally, this would overlay the bubbles’ spectral profiles on top of each other; however, since the redshift upon which each spectrum is centred will be subject to an uncertainty due to the uncertainty in the grism-determined redshift of the target galaxy (which we denote by  $\sigma_z$ ), the bubbles will be scattered along the line-of-sight axis. For the case of GN-z11, Oesch et al. (2016) used *HST*-WFC3 grism spectroscopy in combination with photometric data from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) to place this object at  $z = 11.09^{+0.08}_{-0.12}$ . Even at half of this uncertainty the spatial equivalent of this redshift error at this redshift is approximately twice the size of the bubble surrounding DR-1. Given the design specifications based on the slitless spectroscopic survey capability requirements, Spergel et al. (2013) report that *WFIRST* should be able to determine redshift within  $\sigma_z \leq 0.001(1 + z)$ . We make a conservative assumption in this work by first setting our fiducial redshift error to  $\sigma_z = 0.05$  for all redshifts investigated, but utilize more optimistic values in Section 5 where we explore a method to measure properties of the high-redshift IGM.

##### 4.3.2 Mock observation sets

We simulate the expected redshifted 21-cm stacked spectrum using the *WFIRST*-HLS galaxy survey as follows. First, we only stack redshifted 21-cm spectra corresponding to galaxies brighter than some UV magnitude cutoff,  $M_{\text{UV}}^{\text{cut}}$ , and which have a redshift that falls within some range,  $\Delta z$ , centred on  $z_c$ . Next, we calculate the number of such galaxies in a single SKA1-LOW field using the predicted intrinsic UVLFs provided by Waters et al. (2016) based on the BLUE TIDES simulation<sup>13</sup> (see Appendix C). We find this number to be  $>300$  for the redshifts of interest in this work (see the dashed contour lines in the left-hand and central panels of Fig. 7). We randomly sample these UVLFs to obtain a redshift and magnitude for each galaxy. Then, using our  $\bar{R}-M_{\text{UV}}-z$  model (and estimate for the variance in  $R$ ,  $\sigma_{\bar{R}}^2$ ) described in Section 3.2, together with our chosen value of  $\sigma_z$ , we form a randomly sampled mock observation set consisting of the tuple  $(z, M_{\text{UV}}, R, \Delta d)$  for each galaxy, where  $R$  is the bubble radius (sampled from a Gaussian with mean  $\bar{R}$  and standard deviation  $\sigma_{\bar{R}}$ ), and  $\Delta d$  is the spatial offset of the galaxy from the centre of the spectrum [sampled from a zero-mean Gaussian with standard deviation  $\sigma_d = c\sigma_z/H(z)$ ].

The effective SKA1-LOW field of view has been calculated by applying diffraction theory to a circular aperture and depends on

<sup>11</sup> Note that it is possible to smooth in the sky plane on a scale equal to the diameter of the bubble to reduce noise further with a modest loss of contrast of the bubble. However, due to uncertainty in the grism-determined redshift of the target galaxy and the presence of numerous other bubbles in the surrounding IGM, this tactic does not pay off for imaging *individual* bubbles and makes identifying the bubble difficult, if not impossible.

<sup>12</sup> The centre of the  $\sim 2200$  deg<sup>2</sup> footprint of the *WFIRST*-HLS lies at a declination corresponding to zenith at the future site of the Australian infrastructure for the SKA, the Murchison Radio-astronomy Observatory (<http://www.atnf.csiro.au/projects/askap/site.html>).

<sup>13</sup> We use the results from the BLUE TIDES simulation since its  $(400 h^{-1} \text{ Mpc})^3$  volume includes rarer bright galaxies than *Tiamat*, giving more appropriate UVLF fits.



both the observed wavelength,  $\lambda$ , and station diameter,  $D_s$ . The small angle approximation for this gives

$$\Omega_{\text{SKA}} \approx \frac{\pi}{4} \left( \frac{\lambda}{D_s} \right)^2, \quad (2)$$

providing  $\approx 8 \text{deg}^2$  at 150 MHz, assuming a station diameter of 35 m. Since we find that the number of spectra required for detection is roughly less than half of what we predict is available the field of view is not an active constraint in this work.

### 4.3.3 Other considerations

There are three points to consider before moving on.

(i) Despite the high temporal cadence of our simulation, we have a relatively limited number of snapshots between  $z \sim 9$  and 11. Furthermore, since the *Tiamat* volume is much smaller than the SKA survey volume, the number of bubbles around galaxies available to stack at each redshift is in deficit. This is evident in Fig. 4, which shows only two galaxies (DR-1 and DR-2) brighter than the *WFIRST* detection limit at  $z = 11.1$ . Also, as demonstrated in Section 3, the ionized regions surrounding the brightest galaxies are relatively spherical and isolated (i.e. non-overlapping) at the redshifts investigated in this work ( $z \gtrsim 9$ ). For this reason, we generate synthetic spherical bubbles.

(ii) In order to beat the instrumental noise down to an acceptable level the number of spectra required to be stacked is  $N_{\text{spec}} \gtrsim 50$ . Stack-averaging this number of randomly selected  $\delta T_b$ -fields results in a relatively smooth spectrum (i.e. fluctuations from bubbles in the IGM are averaged out). We therefore make the approximation of embedding our synthetic bubbles in a flat IGM whose globally averaged  $\delta T_b$  is set equation (1).

(iii) In the large- $N_{\text{spec}}$  limit, the stack-averaged signal can be approximated by the convolution of the redshift error probability distribution (mapped on to the space of line-of-sight distance and assumed to be Gaussian) and the spectral profile of a bubble of average size  $R$ , smoothed in the sky plane on the scale of the assumed SKA1-LOW synthesized beam. Depending on the spatial extent corresponding to  $\sigma_z$  ( $\sigma_d$ ) relative to the bubble size, the stack-averaged signal will be quasi-Gaussian (in the case where  $\sigma_d > R$ ), or more closely resemble the individual bubble profile (in the case where  $\sigma_d < R$ ). This is demonstrated in Fig. 8 (see discussion in Section 5.2). We take advantage of this property when calculating ensemble detection statistics in Section 4.3.5.

Taking these issues into consideration, we construct synthetic redshifted 21-cm spectra in the following manner: For each galaxy in the mock observation set with associated properties ( $z$ ,  $M_{\text{UV}}$ ,  $R$ ,  $\Delta d$ ), we embed a spherical bubble of radius  $R$  in a flat IGM volume with  $\delta T_b$  set by equation (1) (using  $\delta = 0$  and  $x_{\text{HI}}$  equal to the interpolated value of the globally averaged neutral fraction at  $z$  based on our fiducial reionization model). Each volume is centred on  $z$  and the bubble is offset from the volume centre by  $\Delta d$ . The brightness temperature field is binned in frequency space and smoothed in the sky plane using a Gaussian beam with a FWHM equal to the average diameter of the bubbles in the set. The ‘image’ in each channel is zero-meaned as interferometers do not measure the DC (zero spacing) mode. The line-of-sight spectrum through the centre of the bubble is then taken and zero-meaned to simulate removal of spectrally smooth extragalactic foregrounds (see the discussion in Section 6). Instrumental noise for each spectrum is simulated by randomly sampling a value for each channel based on  $\sigma_N(\nu, \Delta\nu, \theta_b, t_{\text{int}})$  (noise in each channel is assumed to be uncorrelated). We centre

the noise realization for all individual spectra on the central redshift,  $z_c$ .<sup>14</sup> The stack-average of each of these spectral components is then calculated.

### 4.3.4 Example realizations

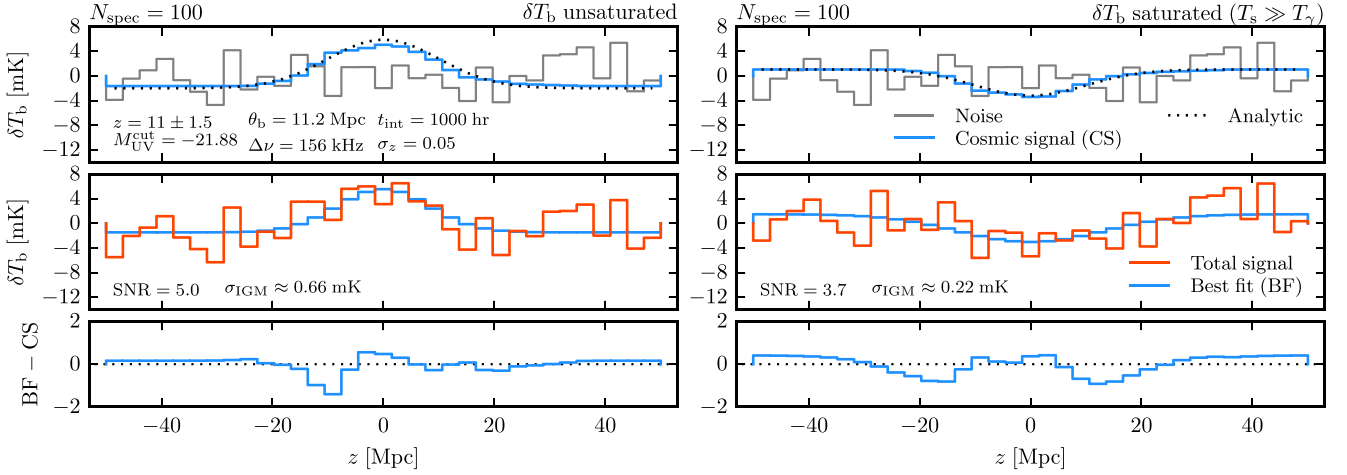
Here we demonstrate our spectral stacking strategy with example realizations. Fig. 6 shows the stack-average of 100 redshifted 21-cm spectra centred on galaxies brighter than  $M_{\text{UV}}^{\text{cut}} = -21.88$  at  $z = 11 \pm 1.5$ . We assume an error in grism-determined redshifts of  $\sigma_z = 0.05$  and a 1000 h integration by SKA. The frequency channel width has been set to 156 kHz (the resulting SNR is insensitive to the choice of reasonable values of  $\Delta\nu$ ). The left-hand panels show the unsaturated case, while the right-hand panels show the saturated case. The upper panels show the two independent components: the cosmic signal (‘CS’, blue) and instrumental noise (grey). The dotted lines are the analytic approximations to the expected stacked cosmic signal. The total signal (cosmic signal + instrumental noise) is shown in the middle panels (red) together with the best-fitting Gaussian model (‘BF’, blue). The Gaussian model is described by two parameters (depth, denoted by  $\Delta T$ , and standard deviation) and has been zero-meaned and fitted with an MCMC parameter estimation technique using flat priors. We calculate the SNR using the resulting marginalized  $\Delta T$  distribution, defining it by  $\text{SNR} = \overline{\Delta T} / \sigma_{\Delta T}$ , where  $\overline{\Delta T}$  is the best-fitting  $\Delta T$  value and  $\sigma_{\Delta T}$  an estimate of the standard deviation derived from the distribution. The SNRs for these example realizations are 5.0 and 3.7 for the unsaturated and saturated cases, respectively. The lower panels show the difference between the input cosmic signal and the best-fitting model (‘BF’ – ‘CS’). Any difference is due to a combination of poor sampling (low  $N_{\text{spec}}$ ), poor fitting and non-Gaussianity. The degree of fluctuation in a randomly stack-averaged IGM for these examples is  $\sigma_{\text{IGM}} \approx 0.66$  mK and 0.22 mK (both  $\ll \overline{\Delta T}$ ) for the unsaturated and saturated cases, respectively.

Naturally, the resulting SNR varies with each realization. For this observation parameter set an ensemble of realizations gives SNRs of  $5.6 \pm 0.9$  and  $3.1 \pm 0.9$  for the unsaturated and saturated cases, respectively. As expected, stack-averaging a larger number of spectra leads to an improvement, e.g. stacking the brightest 300 ( $M_{\text{UV}}^{\text{cut}} = -21.27$ ,  $z = 11 \pm 1.5$ ) gives SNRs of  $7.5 \pm 0.9$  and  $4.6 \pm 0.5$  for the unsaturated and saturated cases, respectively. We now go on to explore the full observational parameter space using ensembles of simulations to gauge detectability, both in terms of the expected average and scatter in the SNR.

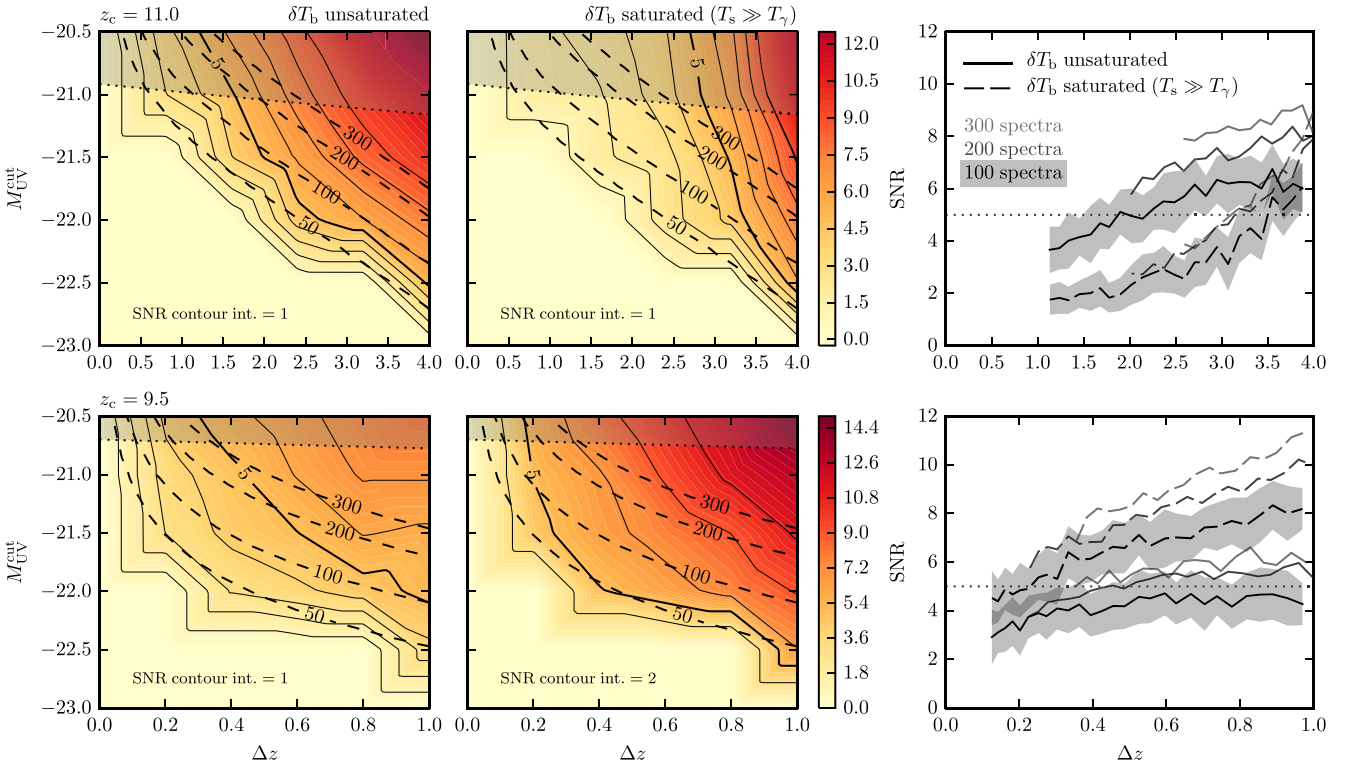
### 4.3.5 Ensemble statistics

The kind of realizations in Section 4.3.4 can be performed anywhere in the valid ( $z_c$ ,  $\Delta z$ ,  $M_{\text{UV}}^{\text{cut}}$ ) observation space. In this section we discuss the average and scatter in SNRs for all the possible observation sets with  $z_c = 9.5$  and 11, for both the unsaturated and saturated signal case. We calculate these by creating an ensemble of 50 realizations at 100 points in the ( $\Delta z$ ,  $M_{\text{UV}}^{\text{cut}}$ ) planes shown in the left-hand and middle column panels of Fig. 7. For computational efficiency, we do this using an analytic formulation of the stack-averaged cosmic signal which we have found to converge on realizations with  $N_{\text{spec}} \gtrsim 50$ . The number of stacked spectra is shown by the dashed line contours, while the colour maps and unbroken

<sup>14</sup> Doing so overestimates the resulting stack-averaged noise since the redshift distribution of the set of mock galaxies is dominated by lower-redshift galaxies (due to the UVLF from which it is drawn).



**Figure 6.** Example stack of (100) redshifted 21-cm spectra centred on galaxies brighter than  $M_{\text{UV}}^{\text{cut}} = -21.88$  at  $z = 11 \pm 1.5$ , assuming an error in grism-determined redshift of  $\sigma_z = 0.05$  and a 1000 h SKA integration. Smoothing in frequency and the image plane has been performed following the guidelines outlined in Sections 4.3.3 and 4.3.4. Left-hand panels show the unsaturated case, while the right-hand panels show the saturated case. Upper panels show the independent components: the cosmic signal (‘CS’, blue) and instrumental noise (grey). Dotted lines are analytic approximations to the expected stacked cosmic signal. The total signal (cosmic signal + instrumental noise) is shown in the middle panels (red) together with the best-fitting Gaussian model (‘BF’, blue). The SNRs for these realizations are 5.0 and 3.7, while the degree of fluctuation in the surrounding stack-averaged IGM is  $\approx 0.66$  mK and 0.22 mK ( $\ll$  signal) for the unsaturated and saturated cases, respectively. Lower panels show the difference between the best-fitting model and the input cosmic signal.



**Figure 7.** Resulting average SNRs for ensembles of realizations exploring the utility of different observation regimes centred on  $z_c = 11$  (upper panels) and  $z_c = 9.5$  (lower panels) for the unsaturated case (left panels) and saturated case (middle panels). Dashed contours show the expected number of galaxies in the SKA field of view brighter than  $M_{\text{UV}}^{\text{cut}}$  with redshifts within  $z_c \pm \Delta z/2$ . The colour maps (and their solid contours) show the estimated average SNR using stacked spectra. The horizontal shaded regions indicate forbidden regimes (dictated by the *WFIRST*  $5\sigma$  sensitivity limit). The right-hand panels show the SNR statistics corresponding to the valid sections of the 100-, 200- and 300-spectra contours as a function of  $\Delta z$ . Solid and dashed lines correspond to the unsaturated and saturated cases, respectively. Shaded regions indicate the  $1\sigma$  confidence regions for the 100-spectra stack results (the variances of the 200- and 300-spectra cases are similar). The dotted line marks a constant SNR of 5.



contours show the average SNR (the bold contours mark a constant SNR of 5, which we use as our threshold for detectability.). The relative two-dimensional gradients of these surfaces at these redshifts suggest that the optimal detection strategy (in terms of maximizing efficiency) is to stack-average bubbles surrounding the brightest galaxies in the widest redshift range possible. The right-hand panels show the SNR statistics corresponding to the valid sections of the 100-, 200- and 300-spectra contours as a function of  $\Delta z$ . Shaded regions indicate the  $1\sigma$  confidence regions for the 100-spectra stack results. Only the 100-spectra scatter is shown as the variances of the 200- and 300-spectra cases are similar. The dotted lines mark our constant SNR = 5 detectability threshold.

## 5 IGM PROPERTIES

### 5.1 Spin temperature

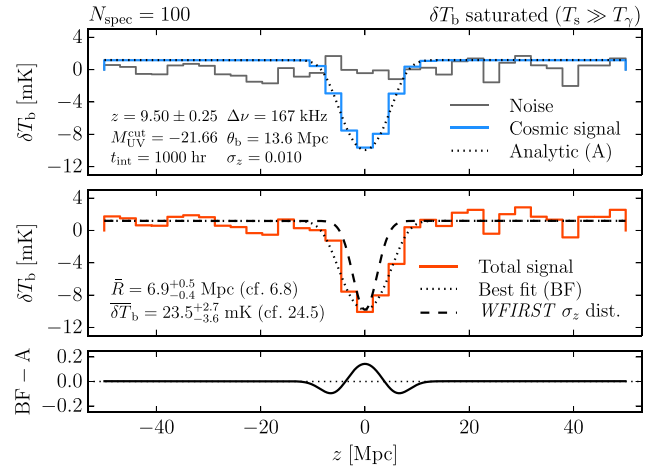
The temperature differential-dependent modality of the 21-cm signal (viz. whether it is in emission or absorption) enables a qualitative constraint to be placed on the average spin temperature of the IGM with respect to the CMB temperature. Given a detection of ionized bubbles using the technique described in the previous section, it is possible to determine if the IGM surrounding them is typically in absorption or emission (see e.g. Fig. 6). If the IGM is, on average, in absorption, then  $T_s < T_\gamma$ . On the other hand, if the IGM is, on average, in emission, then  $T_s > T_\gamma$ .

Furthermore, if the signal mode of the IGM was found to change over a range of redshifts this can be used to provide a quantitative measure of the redshift at which  $T_s = T_\gamma$  and therefore a measure of  $T_s$  at this redshift. This is not possible for the reionization model presented in this work as the bubbles begin to overlap significantly before the IGM begins to appear in emission. However, this may not be the case in reality as heating by unmodelled sources/mechanisms may occur earlier.

### 5.2 Globally averaged neutral fraction and bubble size

Another IGM property of interest is its globally averaged neutral fraction,  $\bar{x}_{\text{HI}}$ . Unfortunately, even if an accurate measurement of the differential brightness temperature in equation (1) is made,<sup>15</sup> this signal depends on both neutral fraction and spin temperature. Therefore, without knowledge of the spin temperature,  $\bar{x}_{\text{HI}}$  can only be determined when the signal from the IGM is saturated ( $1 - T_\gamma/T_s \rightarrow 1$ ). Assuming the IGM is fully heated and the signal appears in saturated emission, we may still be left with a degeneracy in the stacked 21-cm spectra between the average size of the stacked bubbles and the average ionization state of the IGM in which they are embedded. This is because a stack of small bubbles has a similar signature to a stack of larger bubbles in a more ionized IGM (so long as  $\sigma_d > R$ ). This arises due to the grism's limited accuracy and is therefore an observationally introduced degeneracy, not a physical one. We now demonstrate a method which breaks this degeneracy by taking advantage of the non-Gaussianity and/or constant width of the stacked spectral signal observed where uncertainty in the grism-determined redshifts is small (such that  $\sigma_d < R$ ).

<sup>15</sup> This may not be possible using interferometric observations due to the inability to observe the full contrast between fully ionized and not fully ionized regions (due to instrumental resolution, for example), i.e. no zero-signal baseline can be established.



**Figure 8.** Demonstration measuring the average IGM brightness temperature (assumed to be saturated) and bubble size for a mock set of (100) galaxies brighter than  $M_{\text{UV}}^{\text{cut}} = -21.66$  at  $z = 9.5 \pm 0.25$ . A grism-determined redshift error of  $\sigma_z = 0.01$  and a 1000 h SKA integration have been used. The upper panel shows the independent signal components and analytic approximation to the expected stacked cosmic signal ('A', dotted line). The total signal is shown in the middle panel together with the best-fitting analytic model ('BF', dotted line) and *WFIRST*'s grism redshift error distribution (dashed line). The lower panel shows the difference between the best-fitting model and the analytic model.

Using  $\sigma_z \approx 0.001(1 + z)$  as *WFIRST*'s spectroscopic redshift survey capability (Spergel et al. 2013), we have  $\sigma_{z=10} = 0.011 \equiv 2.4$  Mpc. Note that this is less than half of the typical radius of bubbles surrounding galaxies detectable by *WFIRST* at  $z = 10$  according to our simulation. Therefore, rather than being widely spread along the line of sight, the bubbles are relatively tightly aligned on top of each other in the stacked spectrum. As a consequence, their stack-averaged spectrum resembles an instrumentally smoothed bubble profile rather than a Gaussian with a width equal to that of the grism redshift error distribution. The stacked signal can be seen to be in this regime by inspection (in the case of high SNR), or by comparing the width of a simple Gaussian fit for the data to  $\sigma_d$ . Having confirmed  $\sigma_d < R$ , fitting the analytic model for an instrumentally smoothed stack of spectra provides estimates for  $\bar{R}$  and  $\overline{\delta T_b}$ .<sup>16</sup>

This technique is demonstrated for  $z_c = 9.5$  in Fig. 8. The example shown averages 100 redshifted 21-cm spectra centred on galaxies brighter than  $M_{\text{UV}}^{\text{cut}} = -21.66$  with  $z = 9.5 \pm 0.25$ . We assume an error in grism-determined redshift of  $\sigma_z = 0.01$  and a 1000 h integration by SKA. The upper panel shows the cosmic signal (blue) and instrumental noise (grey). The dotted line is the analytic approximation to the expected stacked cosmic signal ('A'). The total signal is shown in the middle panel (red) together with its best-fitting analytic model ('BF', dotted line) and *WFIRST*'s grism redshift error distribution (dashed line, showing that  $\sigma_d < R$ ). Fitting was performed using an MCMC parameter estimation technique using flat priors with an upper limit on  $\overline{\delta T_b}$  equal to the value of  $\delta T_b$  in equation (1) with  $\delta = x_{\text{HI}} = 1$  and assuming  $T_s \gg T_\gamma$ , ensuring an interpreted estimate for  $\bar{x}_{\text{HI}}$  that is physically sensible (i.e.  $\bar{x}_{\text{HI}} \leq 1$ ). The lower panel shows the difference between the best-fitting model and the analytic model ('BF' - 'A'). The resulting best-fitting parameter estimates are  $\bar{R} = 6.9_{-0.4}^{+0.5}$  Mpc (cf. the input value of 6.8 Mpc) and

<sup>16</sup> Note that the  $\overline{\delta T_b}$  estimate is for the average signal of the IGM, not the depth of the trough feature in the observed stacked spectra.

$\overline{\delta T_b} = 23.5_{-3.6}^{+2.7}$  mK (cf. the input value of 24.5 mK). Using equation (1), we estimate (ignoring any covariance between  $\delta T_b$  and  $z$  in error propagation for simplicity) that  $\bar{x}_{\text{H I}} = 0.85 \pm 0.13$  (cf. the input value of 0.89).

In the above example, the observed parameter set used was chosen so as to provide a balance between the size of the bubbles stacked (stacking larger bubbles strengthens the signal thereby improving the quality of the fit) and the redshift range of the galaxies, both of which have an impact on the resulting uncertainty in our interpreted estimate for  $\bar{x}_{\text{H I}}$ .

## 6 DISCUSSION

Some sources embedded in an IGM with  $T_s < T_\gamma$  (i.e. appearing in absorption) will not only ionize their surroundings to form a bubble, but also heat the gas in their proximity through soft X-ray emission. This will give rise to a relatively thin shell of 21-cm emission beyond the bubble (Tozzi et al. 2000; Wyithe & Loeb 2004; Ghara et al. 2016). The  $\delta T_b$ -profile of these sources will therefore resemble a top-hat with ‘horns’. In order to be conservative we have ignored these effects, although they would improve the SNR of the stacked spectra used for our detectability predictions in Section 4.

Foregrounds were anticipated and have proven to be a significant challenge to detecting the cosmic signal due to their brightness and the chromatic response of the new generation of low-frequency interferometers (see e.g. Pober et al. 2016). In this work we have assumed anthropogenic RFI and both Galactic and extragalactic point sources have been removed from the observation data leaving no residual. We have also assumed there are no contamination or removal effects by diffuse Galactic foregrounds (synchrotron emission) apart from the mean removal for each line-of-sight spectra as discussed in Section 4.3.3. Previous work has shown that, due to its spectral smoothness, it is possible to subtract this foreground in the imaging regime using a polynomial fitting-based method along each line of sight (see e.g. McQuinn et al. 2006; Wang et al. 2006; Geil et al. 2008; Petrovic & Oh 2011; Alonso et al. 2015). Since the spectral features we stack in this work are narrow in comparison with the frequency bands over which foreground fitting and subtraction is performed, we expect that these cosmic signatures will not be removed.

## 7 SUMMARY AND CONCLUSIONS

We have investigated the feasibility of directly detecting regions of ionized hydrogen surrounding galaxies by stacking redshifted 21-cm observations around optically identified luminous galaxies during early stages of the EoR. In particular, we look at utilizing low-frequency observations by the SKA and near infrared survey data of *WFIRST* in order to image bubbles surrounding massive galaxies at  $z \gtrsim 9.5$ .

Our main results can be summarized as follows.

(i) Our modelling, using the DRAGONS simulation suite, predicts a linear relationship between the size of ionized bubbles and the luminosity of the brightest galaxy within them (in terms of intrinsic absolute UV magnitude, for  $M_{\text{UV}} \leq -17.25$ ), which evolves with redshift. We provide a fit for this relation and its scatter as a function of redshift.

(ii) Individual bubbles will not be detected with SKA1-LOW. However, by stacking 100 or more redshifted 21-cm spectra it is

possible to detect the EoR directly with a significance of at least  $5\sigma$  at  $z \sim 9-12$ .

(iii) Both the spin temperature of the IGM and the accuracy of the grism-determined redshifts of the galaxies have a significant impact on the detectability of reionization.

(iv) It is possible to place *qualitative* constraints on the evolution of the spin temperature of the IGM at  $z \gtrsim 9$  and it may be possible to *quantitatively* measure the redshift at which it is equal to the CMB temperature.

(v) Measuring the average size of bubbles and globally averaged neutral fraction of the IGM is a difficult task due to the degeneracy of these properties’ contribution to the cosmic signal. However, if the IGM can be assumed to be fully heated (such that the 21-cm signal is saturated) and the accuracy of the grism-determined redshifts of the galaxies is sufficiently high (compared to the average bubble size of the stack) then this degeneracy may be broken and both the average bubble size and neutral fraction can be accurately determined.

We conclude that imaging 21-cm emission around samples of luminous galaxies from the EoR will provide an additional and complementary probe of cosmic reionization.

## ACKNOWLEDGEMENTS

This work was supported by the Victorian Life Sciences Computation Initiative (VLSCI), grant reference UOM0005, on its Peak Computing Facility hosted at The University of Melbourne, an initiative of the Victorian Government. Part of this work was performed on the gSTAR national facility at Swinburne University of Technology. gSTAR is funded by Swinburne University of Technology and the Australian Governments Education Investment Fund. AM acknowledges support from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 638809 AIDA). This research was funded by the Australian Research Council through the ARC Laureate Fellowship FL110100072 awarded to JSBW. Parts of this research was conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020 (PMG). Finally, we would like to thank the anonymous referees for their constructive comments which helped to improve this paper.

## REFERENCES

- Alonso D., Bull P., Ferreira P. G., Santos M. G., 2015, *MNRAS*, 447, 400  
 Bowler R. A. A. et al., 2014, *MNRAS*, 440, 2810  
 Bowman J. D., Rogers A. E. E., Hewitt J. N., 2008, *ApJ*, 676, 1  
 Burns J. O. et al., 2012, *Adv. Space Res.*, 49, 433  
 Cen R., Haiman Z., 2000, *ApJ*, 542, L75  
 Datta K. K., Bharadwaj S., Choudhury T. R., 2007, *MNRAS*, 382, 809  
 Datta K. K., Friedrich M. M., Mellema G., Iliev I. T., Shapiro P. R., 2012, *MNRAS*, 424, 762  
 Fan X., Carilli C. L., Keating B., 2006, *ARA&A*, 44, 415  
 Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181  
 Geil P. M., Wyithe J. S. B., 2008, *MNRAS*, 386, 1683  
 Geil P. M., Wyithe J. S. B., Petrovic N., Oh S. P., 2008, *MNRAS*, 390, 1496  
 Geil P. M., Mutch S. J., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016, *MNRAS*, 462, 804 (Paper V)  
 Ghara R., Choudhury T. R., Datta K. K., 2016, *MNRAS*, 460, 827  
 Ghara R., Choudhury T. R., Datta K. K., Choudhuri S., 2017, *MNRAS*, 464, 2234  
 Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, *MNRAS*, 466, 4239

Kohler K., Gnedin N. Y., Miralda-Escudé J., Shaver P. A., 2005, *ApJ*, 633, 552

Koopmans L. et al., 2015, *Advancing Astrophysics with the Square Kilometre: The Cosmic Dawn and Epoch of Reionisation with SKA*, SISSA, Trieste, PoS(AASKA14), 1

Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2016, *MNRAS*, 462, 235

McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, *ApJ*, 653, 815

Majumdar S., Bharadwaj S., Choudhury T. R., 2012, *MNRAS*, 426, 3178

Mason C. A., Trenti M., Treu T., 2015, *ApJ*, 813, 21

Mesinger A., Furlanetto S., 2007, *ApJ*, 669, 663

Mesinger A., Furlanetto S., Cen R., 2011, *MNRAS*, 411, 955

Mesinger A., Greig B., Sobacchi E., 2016, *MNRAS*, 459, 2342

Morales M. F., Wyithe J. S. B., 2010, *ARA&A*, 48, 127

Mortlock D. J. et al., 2011, *Nature*, 474, 616

Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016a, *MNRAS*, 462, 250 (Paper III)

Mutch S. J. et al., 2016b, *MNRAS*, 463, 3556 (Paper VI)

Oesch P. A. et al., 2016, *ApJ*, 819, 129

Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713

Patra N., Subrahmanyam R., Raghunathan A., Udaya Shankar N., 2013, *Exp. Astron.*, 36, 319

Petrovic N., Oh S. P., 2011, *MNRAS*, 413, 2103

Planck Collaboration XIII, 2016a, *A&A*, 594, A13

Planck Collaboration XLVII, 2016b, *A&A*, 596, A108

Pober J. C. et al., 2016, *ApJ*, 819, 8

Poole G. B., Angel P. W., Mutch S. J., Power C., Duffy A. R., Geil P. M., Mesinger A., Wyithe S. B., 2016, *MNRAS*, 459, 3025 (Paper I)

Salpeter E. E., 1955, *ApJ*, 121, 161

SKAO Science Team 2015, SKA1-low Configuration. Document number: SKA-SCI-LOW-001. Available at: [http://astronomers.skatelescope.org/wp-content/uploads/2015/11/SKA1-Low-Configuration\\_V4a.pdf](http://astronomers.skatelescope.org/wp-content/uploads/2015/11/SKA1-Low-Configuration_V4a.pdf) (SKA2015)

Spergel D. et al., 2013, preprint (arXiv:1305.5422)

Thompson A. R., Moran J. M., Swenson G. W., Jr., 2001, *Interferometry and Synthesis in Radio Astronomy*, 2nd Edn. A Wiley-Interscience publication, Wiley, New York, p. 692

Tozzi P., Madau P., Meiksin A., Rees M. J., 2000, *ApJ*, 528, 597

Wang X., Tegmark M., Santos M. G., Knox L., 2006, *ApJ*, 650, 529

Waters D., Di Matteo T., Feng Y., Wilkins S. M., Croft R. A. C., 2016, *MNRAS*, 463, 3520

Wyithe J. S. B., Loeb A., 2004, *ApJ*, 610, 117

Wyithe J. S. B., Loeb A., Barnes D. G., 2005, *ApJ*, 634, 715

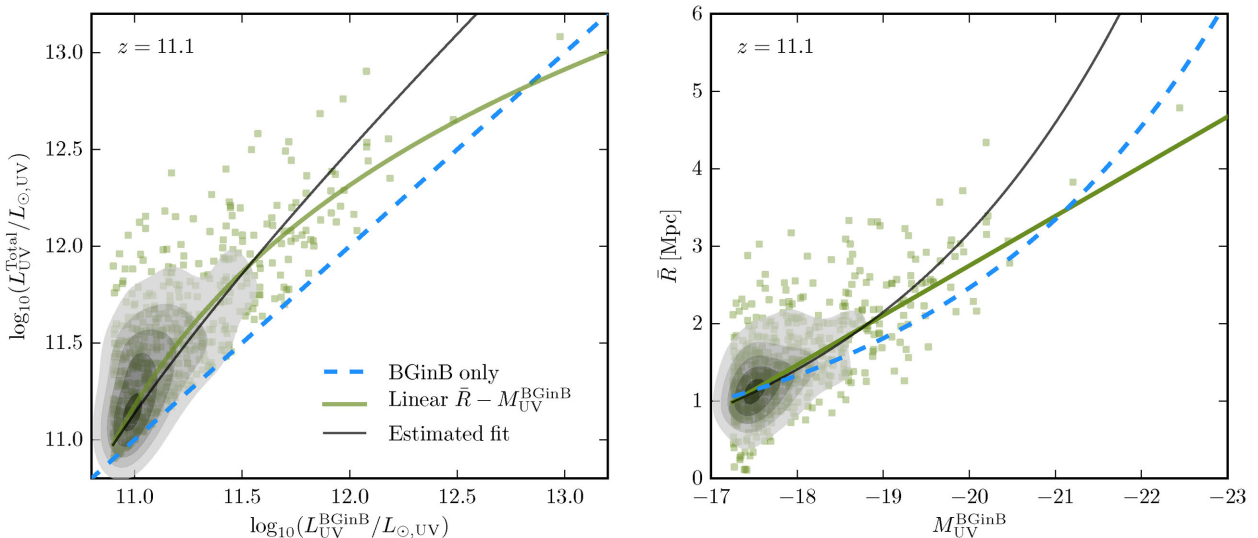
## APPENDIX A: $\bar{R}-M_{\text{UV}}$ RELATION

### A1 Assumed linearity

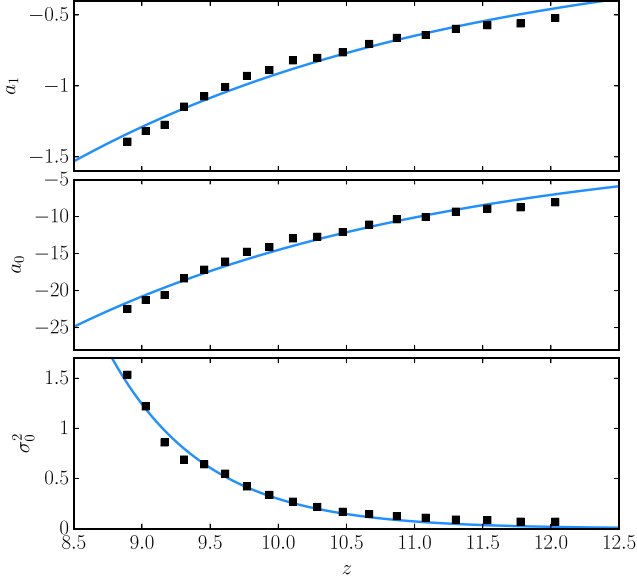
The typical radius of a cosmological Strömgen sphere,  $R_S$ , generated by a source of UV luminosity  $L_{\text{UV}}$  scales as  $R_S \propto L_{\text{UV}}^{1/3}$  (see e.g. Cen & Haiman 2000). In terms of absolute UV magnitude, this gives  $R_S \propto 10^{-0.4M_{\text{UV}}/3}$  (i.e. non-linear in  $R_S-M_{\text{UV}}$ ). The  $\bar{R}-M_{\text{UV}}$  plot in Fig. 4, however, shows average bubble radius as a function of the absolute UV magnitude of the *brightest galaxy only* in each bubble. Of course, many other galaxies contribute ionizing photons towards the formation of an ionized region (the brightest of these can be seen in Figs 1 and 3). These galaxies are both clustered and biased and therefore any enhancement they perform effectively depends on the luminosity of the brightest galaxy in the bubble,  $L_{\text{UV}}^{\text{BGinB}}$ , and the size of the region. The total luminosity of galaxies in a region may therefore be written as

$$L_{\text{UV}}^{\text{Total}} = L_{\text{UV}}^{\text{BGinB}} + L_{\text{UV}}^{\text{NCs}}(L_{\text{UV}}^{\text{BGinB}}, R_S), \quad (\text{A1})$$

(where NCs stands for ‘Non-Centrals’) and the resulting bubble size scales as  $R \propto (L_{\text{UV}}^{\text{Total}})^{1/3}$ . It is this luminosity and scale dependence that alter the simple cube root Strömgen sphere relation. This is clearly demonstrated in the left-hand panel of Fig. A1 which shows  $L_{\text{UV}}^{\text{Total}}$  versus  $L_{\text{UV}}^{\text{BGinB}}$  at  $z = 11.1$  (similar results are obtained at other redshifts.). In this plot the blue dashed line corresponds to the  $L_{\text{UV}}^{\text{Total}} = L_{\text{UV}}^{\text{BGinB}}$  (brightest galaxy-only Strömgen sphere) case, the green line to the linear  $\bar{R}-M_{\text{UV}}$  relation used in this work (normalized to best fit the data), and the thin black line to an estimate of the  $L_{\text{UV}}^{\text{Total}}-L_{\text{UV}}^{\text{BGinB}}$  relation. The important thing to note here is that the data do not suggest a power-law relation between  $L_{\text{UV}}^{\text{Total}}$  and  $L_{\text{UV}}^{\text{BGinB}}$  (and therefore the luminosity enhancement,  $L_{\text{UV}}^{\text{NCs}}$ , is indeed  $L_{\text{UV}}^{\text{BGinB}}$ -dependent). Hence the cube root-in- $L_{\text{UV}}^{\text{BGinB}}$  Strömgen sphere relation cannot hold. The right-hand panel shows the corresponding  $\bar{R}-M_{\text{UV}}^{\text{BGinB}}$  relations (all normalized to best fit the data.). This shows the total luminosity-models (green and black lines)



**Figure A1.** Left panel: total luminosity of galaxies in bubbles as a function of the luminosity of the brightest galaxy in the bubble at  $z = 11.1$ . Right: bubble radius as a function of absolute UV magnitude of the brightest galaxy in the bubble. The blue dashed lines correspond to a brightest galaxy-only Strömgen sphere relationship, the green lines to the linear  $\bar{R}-M_{\text{UV}}$  relation used in this work and the thin black lines to an estimate of the  $L_{\text{UV}}^{\text{Total}}-L_{\text{UV}}^{\text{BGinB}}$  relation.



**Figure A2.** Best-fitting error-weighted  $\bar{R}$ - $M_{\text{UV}}$ -model parameter values (data points) for 18 different redshifts between  $z \sim 9$  and 12:  $a_1$  (upper panel),  $a_0$  (middle panel) and  $\sigma_0^2$  (lower panel). Exponential fits are shown by the blue lines.

**Table A1.** Resulting best-fitting values for the functional coefficients of equation (A2) for each  $\bar{R}$ - $M_{\text{UV}}$  model parameter introduced in Section 3.2.

Parameter description	$\bar{R}$ - $M_{\text{UV}}$ model Symbol	Function coefficients		
		$c_0$	$c_1$	$c_2$
$\bar{R}$ - $M_{\text{UV}}$ gradient	$a_1$	0.0505	0.384	17.5
Systematic $\bar{R}$ offset	$a_0$	2.32	0.401	14.5
Variance of systematic offset	$\sigma_0^2$	2.73e-3	1.37	13.4

appear to fit the data reasonably well (note, however, that the scatter is relatively large.). A reduced chi-squared analysis fails to show that any of these models describe the data significantly better than any other. We leave the analytic solution to the  $\bar{R}$ - $M_{\text{UV}}$  relationship to future work.

### A2 Fitting $\bar{R}$ - $M_{\text{UV}}$ model parameters in redshift

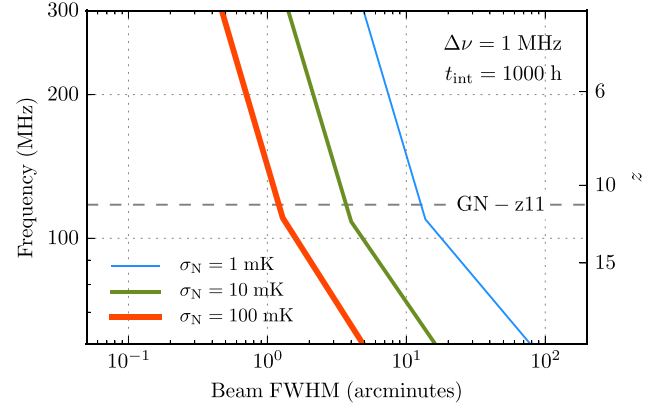
In Section 3.2 we calculated the best-fitting error-weighted  $\bar{R}$ - $M_{\text{UV}}$ -model parameter values for 18 different redshifts between  $z \sim 9$  and 12. These are shown by the data points in Fig. A2. To each parameter ( $a_1$ ,  $a_0$  and  $\sigma_0^2$ ) we fit the same exponential functional form, given by

$$f(z, c_0, c_1, c_2) = c_0 \exp[-c_1(z - c_2)]. \quad (\text{A2})$$

The resulting best-fitting values for these functional coefficients are given in Table A1.

## APPENDIX B: SKA1-LOW NOISE

This Appendix describes how we simulate the instrumental noise of SKA1-LOW. For a comprehensive and authoritative overview of interferometric techniques for radio astronomy, see Thompson, Moran & Swenson (2001).



**Figure B1.** Simulated image-space brightness temperature sensitivity for a deep (1000 h) integration using a frequency channel width of 1 MHz as a function of frequency and synthesized beam FWHM (reproduction of fig. 9 in SKA2015). Blue, green and red lines are the 1, 10 and 100 mK noise isograms, respectively. For reference, the dashed line shows the redshifted 21-cm line frequency corresponding to the redshift of GNz-11.

The image-space noise realizations used throughout this work were generated based on instrumental specifications provided by SKAO Science Team (2015) (hereafter SKA2015). In particular, we use their simulated brightness temperature sensitivity results for a deep (1000 h) integration as a function of frequency and synthesized beam FWHM (see fig. 9 in SKA2015, or the reproduced version shown in Fig. B1). This uses a fiducial frequency channel width of 1 MHz; however, calculating the sensitivity for different integration times and/or frequency channel widths is possible by noting that

$$\sigma_{\text{N}} \propto \frac{1}{\sqrt{\Delta\nu t_{\text{int}}}}. \quad (\text{B1})$$

## APPENDIX C: BLUETIDES UVLF

In this work we have used the following double power law to describe the predicted intrinsic UVLFs for galaxies at  $z = 9$ –11:

$$\phi(M) = \frac{\phi^*}{10^{0.4(\alpha+1)(M-M^*)} + 10^{0.4(\beta+1)(M-M^*)}} \quad (\text{C1})$$

(Bowler et al. 2014). Here  $\phi^*$ ,  $M^*$ ,  $\alpha$  and  $\beta$  are the normalization, characteristic magnitude, faint-end slope and bright end slope, respectively. We use the best-fitting values to these parameters, found using the BLUETIDES simulation by Waters et al. (2016). These are:

$$\ln(\phi^*) = -[(0.96 \pm 0.22)(1+z) + (1.57 \pm 2.32)]$$

$$M^* = (0.28 \pm 0.12)(1+z) + (-24.79 \pm 1.41)$$

$$\alpha = -[(0.14 \pm 0.02)(1+z) + (0.72 \pm 0.24)]$$

$$\beta = -[(0.15 \pm 0.05)(1+z) + (1.78 \pm 0.60)]$$

where  $\phi^*$  is in  $\text{Mpc}^{-3} \text{mag}^{-1}$ .

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