Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **483,** 1980–1992 (2019)
Advance Access publication 2018 November 30

doi:10.1093/mnras/sty3260

What does the first highly redshifted 21-cm detection tell us about early galaxies?

Jordan Mirocha ★ and Steven R. Furlanetto
Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA

Accepted 2018 November 28. Received 2018 November 28; in original form 2018 March 8

ABSTRACT

The Experiment to Detect the Global Epoch of Reionization Signature (EDGES) recently reported a strong 21-cm absorption signal relative to the cosmic microwave background at $z \sim 18$. While its anomalous amplitude may indicate new physics, in this work we focus on the timing of the signal, as it alone provides an important constraint on galaxy formation models. Although rest-frame ultraviolet luminosity functions (UVLFs) over a broad range of redshifts are well fit by simple models in which galaxy star formation histories track the assembly of dark matter haloes, we find that these same models, with reasonable assumptions about X-ray production in star-forming galaxies, cannot generate a narrow absorption trough at $z \sim 18$. If verified, the EDGES signal therefore requires the fundamental inputs of galaxy formation models to evolve rapidly at $z \gtrsim 10$. Unless extremely faint sources residing in haloes below the atomic cooling threshold are responsible for the EDGES signal, star formation in $\sim 10^8$–10^{10} $M_\odot$ haloes must be more efficient than expected, implying that the faint end of the UVLF at $M_{\rm UV} \lesssim -12$ must *steepen* at the highest redshifts. This steepening provides a concrete test for future galaxy surveys with the *James Webb Space Telescope* and ongoing efforts in lensed fields, and is required regardless of whether the amplitude of the EDGES signal is due to new cooling channels or a strong radio background in the early Universe. However, the radio background solution requires that galaxies at $z > 15$ emit 1–2 GHz photons with an efficiency $\gtrsim 10^2$ times greater than local star-forming galaxies, posing a challenge for models of low-frequency photon production in the early Universe.

Key words: galaxies: high-redshift – intergalactic medium – galaxies: luminosity function, mass function – dark ages, reionization, first stars – diffuse radiation.

1 INTRODUCTION

A simple picture of high-z galaxy evolution has begun to emerge in recent years as the rest-frame ultraviolet luminosity function (UVLF) of galaxies has been assembled by deep surveys at $4 \lesssim z \lesssim 10$ (e.g. Bouwens et al. 2015b; Finkelstein et al. 2015). The rapid decline in the abundance of bright galaxies (e.g. Oesch et al. 2018), coupled with a reduction in the measured value of the Thomson scattering optical depth of the cosmic microwave background (Planck Collaboration 2016), has reduced the need for an abundance of galaxies far beyond detection thresholds at $z \gtrsim 6$. If one extrapolates UVLFs to absolute magnitudes $M_{\rm UV} \gtrsim -12$ (roughly the atomic cooling threshold in most models), reionization ends at $z \sim 6$, so long as the escape fraction of UV photons is $0.1 \lesssim f_{\rm esc} \lesssim 0.2$ (e.g. Bouwens et al. 2015a; Robertson et al. 2015).

21-cm observations have long been expected to provide a complementary view of structure formation in the early Universe, and thus an independent check on this emerging picture (Madau, Meiksin & Rees 1997). Fluctuations in the background, being targeted by interferometers like PAPER (Parsons et al. 2014), MWA (Dillon et al. 2014), LOFAR (Patil et al. 2017), and HERA (DeBoer et al. 2017), trace spatial variations in the density, ionization (Furlanetto, Zaldarriaga & Hernquist 2004), kinetic temperature (Pritchard & Furlanetto 2007), and Ly-α (Barkana & Loeb 2005) field, and can thus constrain the spatial distribution of sources and their radiative properties. Alternatively, the sky-averaged ('global') 21-cm signal (Shaver et al. 1999) traces the volume-averaged ionization and thermal histories in time, and thus offers a powerful independent constraint on the timing of reionization and reheating.

Though early work predicted a global signal with only a weak high-frequency emission feature, subsequent papers found a much richer structure (e.g. Furlanetto 2006; Pritchard & Loeb 2010; Mesinger, Ferrara & Spiegel 2013) due to the inefficiency of Ly-α heating (Chen & Miralda-Escudé 2004; Chuzhoy, Alvarez &

★ E-mail: mirocha@astro.ucla.edu

© 2018 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

Shapiro 2006; Furlanetto & Pritchard 2006; Hirata 2006), which results in an intergalactic medium (IGM) that remains cooler than the cosmic microwave background (CMB) prior to (and in some cases during) reionization. Indeed, some theoretical models in recent years predict a relatively cold IGM during reionization, if the dominant X-ray sources in high-z star-forming galaxies are high-mass X-ray binaries (HMXBs, as is the case locally; Gilfanov, Grimm & Sunyaev 2004; Mineo, Gilfanov & Sunyaev 2012), since HMXBs have hard spectra (Fialkov, Barkana & Visbal 2014; Mirocha 2014), and because of the apparent inefficiency of star formation in faint galaxies, which delays the onset of heating (Mesinger, Greig & Sobacchi 2016; Madau & Fragos 2017; Mirocha, Furlanetto & Sun 2017). Population III stars and proto-quasars likely have a more subtle impact on the signal (Tanaka, O'Leary & Perna 2016; Mirocha et al. 2018). For a review of high-z galaxy formation and 21-cm cosmology, see e.g. Barkana & Loeb (2001), Furlanetto, Oh & Briggs (2006), and Pritchard & Loeb (2012).

The first reported detection from the Experiment to Detect the Global Epoch of Reionization Signature (EDGES; Bowman et al. 2018) is consistent with a cold IGM before reionization[1] – in fact, even colder than that expected in an adiabatically cooling IGM.

The amplitude of the EDGES signal could be evidence of interactions between baryons and dark matter (Barkana 2018), though a weakly charged dark matter (DM) particle (capable of cooling the baryons through Rutherford scattering) cannot account for the EDGES signal without causing tension elsewhere. For example, Muñoz & Loeb (2018) estimate that if there is a charged DM particle, it can only constitute ∼10 per cent or less of all of the DM.

Alternatively, Feng & Holder (2018) showed that a high-z radio background could supplement the CMB as the illuminating backdrop against which 21-cm absorption is measured, and thus bias the global 21-cm signal towards larger amplitudes. In this work, we will explore both scenarios via flexible additions to our baseline models.

In Mirocha et al. (2017, hereafter M17), we put forth a set of predictions for the global 21-cm signal calibrated to high-z UVLF measurements and with empirically motivated choices for X-ray source populations. The generic result was a strong absorption trough at high frequencies, $\nu \sim 100 \pm 10$ MHz. Some models of this variety were quickly disfavoured, both by EDGES (Monsalve et al. 2017) and SARAS (Singh et al. 2017), though most remained viable until the recent report from EDGES (Bowman et al. 2018). A ∼78 MHz trough, regardless of its amplitude, cannot be accommodated by these models in their current form.

Because the M17 models were constructed assuming current knowledge of stellar populations and physically motivated extrapolations of the LF, deviations from our predictions suggest that the UVLF-based view of high-z galaxy evolution is incomplete, regardless of any exotic processes that amplify the signal's strength. A natural first supposition is that very faint sources, like Population III stars and their remnants, could drive the signal to lower frequencies. However, such sources only qualitatively modify the ∼100 MHz expectations under extreme circumstances (Mebane, Mirocha & Furlanetto 2018; Mirocha et al. 2018). Given the tension between these predictions and the recent EDGES measurement, our goal in this work is to determine what aspects of the astrophysical model – which is representative of the kinds of models used broadly in the high-z galaxy evolution community (e.g. Dayal et al. 2014; Behroozi & Silk 2015; Mason, Trenti & Treu 2015; Mashian, Oesch & Loeb 2016; Sun & Furlanetto 2016) – require revision.

In order to test the robustness of our conclusions about high-z galaxies, we explore both signal-amplifying mechanisms (i.e. Barkana 2018; Feng & Holder 2018) using flexible parametric models. That is, rather than predicting the source(s) responsible for the stronger-than-expected EDGES signal from physical arguments, we attempt to infer the thermal history and radio background that are required to explain the data. Reassuringly, we find that our inferences about galaxy evolution are largely insensitive to the method by which one amplifies the signal's strength. However, the radio background approach places extreme requirements on low-frequency photon production in star-forming galaxies, as we will discuss in Section 3.

We introduce the 21-cm signal and review our galaxy evolution modelling procedure in Section 2, and present results, discussion, and conclusions in Sections 3, 4, and 5, respectively. We adopt Planck Collaboration (2016) cosmological parameters throughout.

2 METHODS

In this section, we briefly review the global 21-cm signal (Section 2.1), the main features of our galaxy evolution model (Section 2.2), and two possible solutions to the signal's anomalous amplitude (Section 2.3). The core components of the model (Sections 2.1–2.2) are simple extensions of those described in Section 2 of M17, so readers familiar with this model may skip ahead to Section 2.3.

2.1 The global 21-cm signal

We employ a simple model for the global 21-cm signal in which the IGM is partitioned into two phases (e.g. Furlanetto 2006; Pritchard & Loeb 2010). The first is a fully ionized phase, whose sole characteristic is its volume-filling factor, $Q_{\rm H\,II}$. The second component of the IGM is often referred to as the 'bulk' IGM, i.e. the IGM outside of ionized bubbles, whose temperature, $T_{\rm K}$, electron fraction, $x_{\rm e}$, and Ly-α intensity, J_α, must all be considered, as they govern the excitation (or 'spin') temperature of HI, $T_{\rm S}$. The global 21-cm signal is simply the volume-averaged brightness temperature of the bulk IGM, down-weighted by the fraction of the volume that is ionized, measured relative to the radiation background temperature,[2] usually the CMB, i.e.

$$\delta T_{\rm b} \simeq 27(1-\overline{x}_i)\left(\frac{\Omega_{\rm b,0}h^2}{0.023}\right)\left(\frac{0.15}{\Omega_{\rm m,0}h^2}\frac{1+z}{10}\right)^{1/2}\left(1-\frac{T_{\rm R}}{T_{\rm S}}\right), \quad (1)$$

where $\overline{x}_i = Q_{\rm H\,II} + (1-Q_{\rm H\,II})x_{\rm e}$ is the volume-averaged ionized fraction, and

$$T_{\rm S}^{-1} \approx \frac{T_{\rm R}^{-1} + x_{\rm c}T_{\rm K}^{-1} + x_\alpha T_\alpha^{-1}}{1 + x_{\rm c} + x_\alpha}. \quad (2)$$

In words, equation (2) means that the spin temperature of HI is set by collisions, whose coupling strength depends on the coefficient $x_{\rm c}$ (which itself depends on density, temperature, and ionization state; Zygelman 2005) and the kinetic temperature (determined by the balance between heating from sources and cooling due to cosmic expansion), and radiation backgrounds, including the sum of

[1] This result is consistent with previous limits from PAPER (Parsons et al. 2014; Pober et al. 2015), which required only that the IGM have been heated prior to $z \sim 9$.

[2] The most commonly used expression for this so-called 'differential brightness temperature' is a just solution to the 1D radiative transfer problem, having replaced the usual intensities with temperatures, i.e. adopting the Rayleigh–Jeans approximation).

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

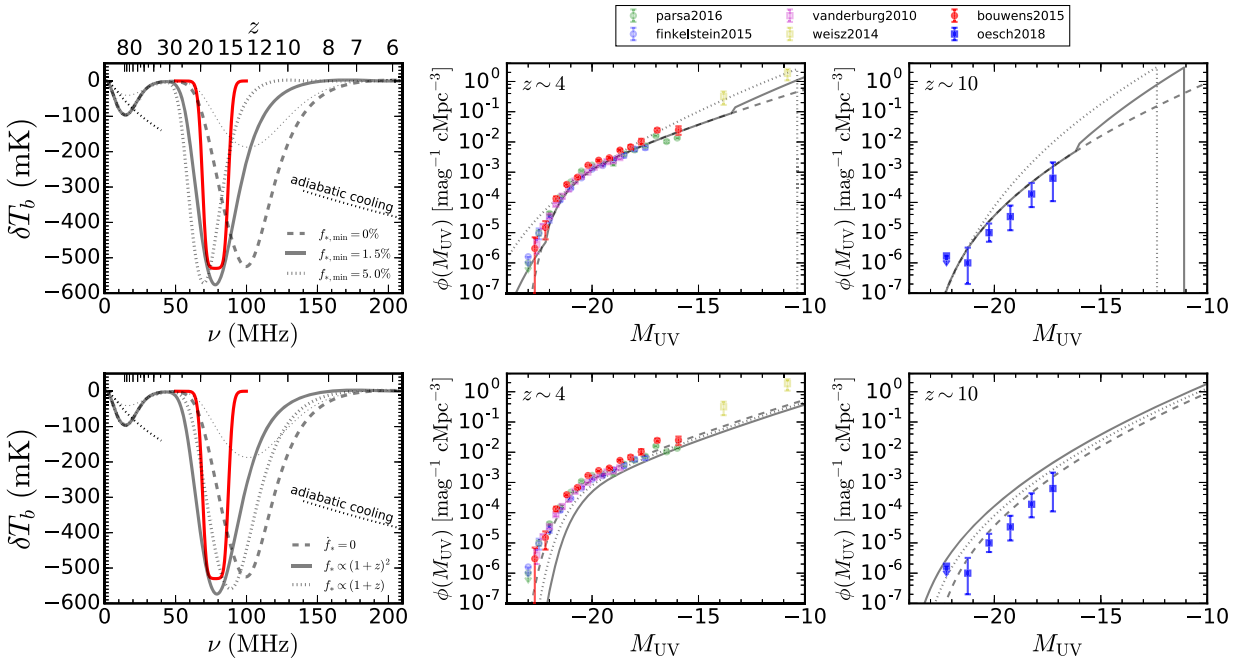

Figure 1. Competing requirements of UVLFs and the EDGES signal. In general, driving the 21-cm absorption trough requires the addition of 'new' star formation – the top row appeals to a change in the shape of the SFE in low-mass objects (by introducing a floor in the SFE), while the bottom row instead invokes time evolution in the overall normalization of the SFE. *Left:* Both methods of boosting star formation at high z can recover 21-cm absorption troughs at ∼80 MHz, though the modifications required to do so are substantial, i.e. a 1.5 per cent floor in the SFE (top left) or $(1+z)^2$ evolution in its normalization (bottom left). However, our example SFE modifications manifest in the UVLF as well, as shown at $z \sim 4$ (*middle column*) and $z \sim 10$ (*right column*). Enhanced star formation in low-mass objects steepens the faint end of the UVLF, which, if too aggressive, causes tension with directly measured high-z UVLFs. However, the UVLF reconstructions from Weisz, Johnson & Conroy (2014) at $z \sim 4$ prefer a steepening. Redshift evolution in the SFE normalization (bottom) leads to considerable tension with both the $z \sim 4$ and $z \sim 10$ LF constraints, for models that match at $z \sim 6$ by construction, assuming the normalization evolves monotonically in time as a power law (indicated in legend). Note that if $z \sim 10$ galaxies have systematically low dust contents, the disagreement at $z \sim 10$ will become even worse.

CMB and any excess in the Rayleigh–Jeans tail, with temperature $T_{\rm R} = T_\gamma + T_{\rm excess}$, and the Ly-$\alpha$ background, characterized by T_α. The latter dependence is not obvious – spontaneous absorption and re-emission of Ly-α photons can result in a spin-flip (see e.g. Fig. 1 of Pritchard & Furlanetto 2006), a byproduct of quantum selection rules for the total spin angular momentum. This is known as the Wouthuysen–Field effect (Wouthuysen 1952; Field 1958), with a magnitude quantified by the radiative coupling coefficient is $x_\alpha = 1.81 \times 10^{11} \widehat{J}_\alpha S_\alpha / (1+z)$, where J_α is the background Ly-α intensity. The factor S_α accounts for line profile effects (Chen & Miralda-Escudé 2004; Chuzhoy et al. 2006; Furlanetto & Pritchard 2006; Hirata 2006). The high optical depth of the Ly-α line rapidly equilibrates the radiation and kinetic temperatures such that in practice, $T_\alpha \simeq T_{\rm K}$.

The time evolution of the IGM's properties, $Q_{\rm H\,II}$, $T_{\rm K}$, and $x_{\rm e}$, is of course governed by the properties of sources in the volume. We construct a model for the volume-averaged emissivity of sources as a function of time (to be discussed in the next subsection), and evolve the emergent radiation field using standard techniques (as in, e.g. Haardt & Madau 1996) to obtain the mean meta-galactic background intensity, J_ν, at all redshifts. With J_ν in hand, ionization and heating rates can be calculated, and the state of the gas in each phase of our two-zone IGM can be evolved in time. We perform these computations with the ARES code,[3] which has been described in greater detail elsewhere (Mirocha 2014).

[3] https://bitbucket.org/mirochaj/ares

2.2 High-z galaxies

We adopt a simple model for high-z galaxy evolution, which assumes that star formation is fuelled by the smooth inflow of pristine gas from the IGM into galaxies, i.e. $\dot{M}_* = f_* \dot{M}_{\rm h}$. Similar techniques have been employed by several groups in recent years (e.g. Behroozi & Silk 2015; Mason et al. 2015; Mashian et al. 2016; Sun & Furlanetto 2016). Although some models use merger trees to construct halo growth trajectories, we adopt a simpler approach. The growth rate of haloes, $\dot{M}_{\rm h}$, is computed assuming haloes grow at fixed number density. Though an oversimplification, this yields fair agreement with the results of numerical simulations (e.g. McBride, Fakhouri & Ma 2009) where they overlap (at $3 \lesssim z \lesssim 6$), which find roughly that $\dot{M}_{\rm h} \propto M(1+z)^{5/2}$. This approach is convenient also because it guarantees self-consistency with the adopted halo mass function[4] (see Furlanetto et al. 2017, Appendix A).

With a model for the abundance of galaxies and their growth rates, one can empirically calibrate the star formation efficiency (SFE), f_*, by fitting the model to UVLF measurements. We take the SFE to be a double power law in halo mass,

$$f_*(M_{\rm h}) = \frac{f_{*,0}}{\left(\frac{M_{\rm h}}{M_{\rm p}}\right)^{\gamma_{\rm lo}} + \left(\frac{M_{\rm h}}{M_{\rm p}}\right)^{\gamma_{\rm hi}}}, \quad (3)$$

[4] We assume a Sheth, Mo & Tormen (2001) form, which we generate using the HMF code (Murray, Power & Robotham 2013).

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

where $f_{*,0}$ is the SFE at its peak mass, M_p, and γ_{lo}, and γ_{hi} describe the power-law index at low and high masses, respectively. We allow further complexity at the faint end through either (i) a 'floor' in the SFE, imposed simply by taking the maximum value of a free parameter $f_{*,min}$ and that of equation (3), or (ii) a steep decline in the SFE, designed to mimic the fraction of haloes occupied by actively star-forming galaxies in the simulations of O'Shea et al. (2015),

$$f_{*,occ} = \left[1 + \left(2^{\mu/3} - 1\right)\left(\frac{M_h}{M_c}\right)^{-\mu}\right]^{-3/\mu}. \quad (4)$$

We explored both cases in Mirocha et al. (2017), but will find here that only the former model can alleviate tension between UVLFs and the EDGES signal.

We assume that the luminosity of galaxies at all wavelengths is dominated by star formation, such that $L_{h,\nu} = \dot{M}_*(M_h, z)l_\nu$, where l_ν is the specific luminosity per unit star formation rate (SFR). Then, assuming a 1:1 correspondence between DM haloes and galaxies, we can derive the UVLF (at the usual rest-frame 1600 Å) as

$$d\phi(L_h) = \frac{dn(M_h, z)}{dM_h}\left(\frac{dL_h}{dM_h}\right)^{-1} dL_h, \quad (5)$$

where $n(M_h, z)$ is the number density of haloes of mass M_h at redshift z, and ϕ is the number density of galaxies with luminosity L_h.

In M17, we calibrated to the Bouwens et al. (2015b) UVLF at $z = 5.9$ assuming a double power-law form for the SFE. Though a redshift-independent SFE is known to result in galaxy populations broadly consistent with current measurements over $0 \lesssim z \lesssim 8$ (e.g. Behroozi, Wechsler & Conroy 2013; Mason et al. 2015), application of these models to very high z cannot explain the EDGES measurement, as we will show in the next section. As a result, we allow the normalization, peak mass, and slopes of the SFE (at masses above and below the peak) to evolve with redshift as power laws. In addition, we allow the SFE to deviate from a single power law at the faint end, either by reaching a floor, or declining exponentially to zero below some critical mass. These phenomenological extensions are allowed to vary with redshift as well. We do not attempt to model scatter in the luminosity of galaxies (at fixed mass), caused either by scatter in the dust correction or SFRs, as these effects mostly affect the bright end of the luminosity function, to which the global 21-cm signal is least sensitive.

The volume-averaged emissivity, which seeds the meta-galactic radiation background, can be computed as a weighted integral over the luminosity function assuming that radiation at all wavelengths is dominated by star formation. We use the BPASS version 1.0 models (Eldridge & Stanway 2009) to generate the UV luminosities of galaxies assuming continuous star formation, and adopt an empirically calibrated relation between SFR and X-ray luminosity, L_X, to synthesize the X-ray background. Our fiducial model assumes the L_X-SFR relation from Mineo et al. (2012), who find that HMXBs are the dominant sources of X-ray emission in local star-forming galaxies, emitting $L_X = 2.6 \times 10^{39}$ erg s^{-1} ($M_\odot$/yr)$^{-1}$ in the 0.5–8 keV band. We assume an unabsorbed multicolour disc spectrum (Mitsuda et al. 1984) extending from 0.2 keV to 30 keV for our calculations. Order of magnitude boosts in the L_X–SFR relation may be possible at high z in the low-metallicity environments (Brorby et al. 2016) of the first galaxies, so we allow the L_X-SFR normalization to scale by a factor, f_X, left as a free parameter in our fits.

In this work, we assume that only galaxies in haloes above the atomic cooling threshold, with virial temperatures $T_{min} \gtrsim 10^4$ K, contribute to the volume-averaged emissivity. Our goal is to determine if the known galaxy population, with some extrapolation to magnitudes and redshifts beyond current detection thresholds, can account for the EDGES signal. Entirely 'new' sources (i.e. those that have yet to be observed directly), of which Pop III stars and their remnants are plausible candidates, could of course also impact the signal. However, Pop III stars should have a rather subtle impact on the global 21-cm signal in all but the most extreme cases (Mirocha et al. 2018),[5] hence our focus on more massive galaxies for the time being. There could be some ambiguity in the interpretation of our results, even if Pop III stars are not ultimately important. For example, are the extrapolations we find necessary to accommodate the EDGES signal indicative of real changes in how galaxies function at the highest redshifts? Or, are they symptoms of problems with the model itself, or neglect of other, more familiar source populations (e.g. globular clusters, active galactic nuclei)? We defer a detailed discussion of this issue to Section 4.

Finally, before moving on, we emphasize that equation (5) is equivalent to the *observed* LF only under the assumption that all photons in the observed band (rest-frame 1600 Å here) escape galaxies. In general, this is not the case, as some rest-frame 1600 Å photons will be absorbed by dust before they can escape the galaxy. As a result, we extend the M17 calibration to the entire redshift range $4 \lesssim z \lesssim 10$ using a Meurer, Heckman & Calzetti (1999) relation between extinction, A_{UV}, and UV slope, β, inferring β through the β–M_{UV} relation of (Bouwens et al. 2014). This is in contrast to M17, in which we neglected dust, in part motivated by the seemingly low dust content in some high-z galaxies (Capak et al. 2015). Including dust has only a minor effect on the global signal – changing the star formation rate density (SFRD) by a factor of ∼2 – though the impact on UVLFs is large, especially at the bright end. Rising dust temperatures could be responsible for the observed dust deficit (e.g. Narayanan et al. 2018), though we make no attempt to model the time evolution of the dust contents of galaxies in this work (as in, e.g. Imara et al. 2018).

2.3 Mechanisms for amplifying the differential brightness temperature

In order to even qualitatively match the EDGES measurement our models require a process capable of increasing the amplitude of the global 21-cm absorption signal beyond that which is expected in an adiabatically cooling IGM. This can be accomplished via new coolants, such as charged DM (Barkana 2018), which can reduce the gas temperature below that predicted in standard lambda cold dark matter models of recombination (e.g. using COSMOREC; Chluba & Thomas 2011). Alternatively, any radio backgrounds present with an intensity comparable to the CMB would supplement T_γ in equation (1) and cause stronger absorption signals (at fixed T_S). Indeed, such backgrounds could amplify the global 21-cm signal substantially without causing tension elsewhere (Feng & Holder 2018), as there is an excess in the observed cosmic radio background at frequencies $\nu \lesssim 1$ GHz of order ∼few Kelvin, rapidly rising to $\gtrsim 10^3$ K at $\nu \lesssim 100$ MHz (Fixsen et al. 2011). The observed excess could be a result of instrumental systematics or new sources (see conference summary by Singal et al. 2018), hence the need to explore the plausibility of new sources at high z in light of the strong EDGES signal.

[5] In Mirocha et al. (2018), we deemed 'extreme' those models in which all H_2-cooling minihaloes host Pop III star formation and all Pop III stars are massive ∼100 $M_\odot$ objects that form in clusters.

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

We explore the effects of both mechanisms for global 21-cm signal amplification, which we describe briefly below.

2.3.1 A parametric approach to 'excess' cooling

Barkana (2018) found that millicharged DM can provide an additional cooling channel for the baryons so long as its mass is $m_\chi < 23$ GeV and its cross-section is $\sigma > 3.4 \times 10^{-21}\,\mathrm{cm}^2$. However, Muñoz & Loeb (2018) have argued that this type of DM cannot constitute the entirety of DM without violating constraints on the local DM density. Given that the origin of the excess cooling is still up for debate, we take a model-agnostic view. Rather than appealing to a particular physical model for DM, we parametrize the thermal history using a form that can (i) accurately recover a case in which the thermal history proceeds 'normally,' i.e. as recombination codes predict (e.g. COSMOREC; Chluba & Thomas 2011), and (ii) allow cooling to occur more rapidly and/or at earlier times than in typical models.

To do this, we recognize that at high z the temperature of a mean-density gas parcel evolves between $T(z) \propto (1+z)$ and $T(z) \propto (1+z)^2$ – the $(1+z)$ dependence a signature that Compton scattering tightly couples the CMB, spin, and kinetic temperatures, and the $(1+z)^2$ dependence indicating that Compton scattering has become inefficient, allowing the gas to cool adiabatically. The simplest parametrization of the thermal history thus appears to be a broken power-law in redshift. However, such a thermal history can also be recovered by noting that the *log*-cooling rate, $d\log T/d\log t$, of a mean-density gas parcel transitions smoothly between $-2/3$ at very high z and $-4/3$ in a matter-dominated cosmology. So, rather than modelling the thermal history directly, we take

$$\frac{\mathrm{d}\log T}{\mathrm{d}\log t} = \frac{\alpha}{3} - \frac{(2+\alpha)}{3}\left\{1 + \exp\left[-\left(\frac{z}{z_0}\right)^{\beta}\right]\right\} \tag{6}$$

and integrate to obtain the thermal history. We have constructed this relation such that $\alpha = -4$ reproduces the typical thermal history, and while varying α can change the late-time cooling rate, the cooling rate as $z \to \infty$ tends to $d\log T/d\log t = -2/3$, as it is must to preserve the thermal history during the recombination epoch. We find that this form does a better job recovering thermal histories from COSMOREC than does a broken power law for $T(z)$. It also outperforms a hyperbolic tanh function for the log cooling rate, which is another natural choice when modelling smooth transitions between known limits. Setting $z_0 = 189.6$, $\beta = 1.27$, and $\alpha = -4$ reproduce the thermal history and its derivative generated with COSMOREC [assuming Planck Collaboration et al. (2016) cosmological parameters] to better than 1 per cent. The parameter z_0 is a decoupling redshift, indicating the redshift at which the cooling rate is halfway between the asymptotic limits, while β indicates how rapidly the cooling rate declines after Compton scattering becomes inefficient. Larger values of β indicate sharper declines in $d\log T_\mathrm{K}/d\log t$. There is of course a degeneracy between these parameters, but, for the purposes of interpreting the $z \sim 18$ EDGES measurement, the detailed values of z_0 and β are unimportant. Measuring the underlying cooling rate will require observations at $\nu \lesssim 20$ MHz, where modifications to standard thermal histories can be studied without the complicating influence of astrophysical sources (see fig. 1 of Fialkov, Barkana & Cohen 2018). Such observations will likely require observations from space (e.g. DARE; Burns et al. 2017).

We allow astrophysical sources to begin forming at $z = 60$, at which point we solve the thermal history by numerically integrating the standard evolution equations, which depend on the rate of heating (from sources) and cooling (from equation 6). Note that while we do not self-consistently solve for the very high-z thermal and ionization histories, any modifications to the post-recombination electron fraction have very little impact on the 21-cm background, since H-e^- collisions are subdominant to H–H collisions in setting the spin temperature unless the electron fraction is order unity.

2.3.2 A radio background excess

The global 21-cm signal may also be amplified relative to common expectations if the CMB is not the only radio background at very high z. Feng & Holder (2018) have worked backwards from the ARCADE-2 excess (Fixsen et al. 2011), and found that stronger-than-expected global 21-cm signals are possible even if only ∼10 per cent of the $z = 0$ excess is from high z. However, they did not attempt to model the source of such a high-z background.

We investigate the possibility that star-forming galaxies themselves source the radio background, and thus account for all features of the observed global 21-cm signal. Accreting black holes could also give rise to such emission (Ewall-Wice et al. 2018), and though we do not include their contribution explicitly here, our calculations do place limits on any scenario in which black hole growth and star formation are tightly related.[6]

There are known relations between SFR and radio luminosity (e.g. Condon, Cotton & Broderick 2002; Heesen et al. 2014; Gürkan et al. 2018), which we can easily implement in our models and augment T_γ accordingly at all redshifts. As our reference point, we adopt an empirical relation between the monochromatic 150 MHz luminosity and SFR (as in, e.g. Gürkan et al. 2018),

$$L_\mathrm{R} = 10^{22} f_\mathrm{R} \left(\frac{\mathrm{SFR}}{\mathrm{M}_\odot\ \mathrm{yr}^{-1}}\right)\ \mathrm{W\ s^{-1}\ Hz^{-1}} \tag{7}$$

and extrapolate to higher frequency assuming a spectral index of -0.7, as in Gürkan et al. (2018).

We treat f_R as a free parameter, and aim to quantify what its value must be in order to explain the EDGES measurement. We do not allow the spectrum to evolve, though in reality, it is likely to steepen at very high z as inverse-Compton losses become increasingly important (Oh 2001). This effect will bias our estimates of f_R low (perhaps by a factor of $\gtrsim 10^3$; Sharma 2018). Also, we will show in the next section that truncating the radio background at some critical redshift, z_off, is necessary in order to match both the shape of the EDGES signal and to satisfy the $z = 0$ ARCADE-2 excess. Consideration of a more exotic source of (non-astrophysical) low-frequency photons may thus be warranted (e.g. Fraser et al. 2018; Pospelov et al. 2018).

All available parameters are summarized in Table 1.

3 RESULTS

3.1 Qualitative expectations

The absorption peak detection by EDGES is centred at $\nu \sim 78$ MHz, some ∼10–30 MHZ lower in frequency than the reference models from M17. In other words, Wouthuysen–Field coupling and X-ray

[6]We assume the background is spatially uniform, though fluctuations in the radio background can have interesting signatures as well (e.g. Ewall-Wice et al. 2014; Bolgar et al. 2018).

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

Table 1. Complete listing of available parameters (left), including their functional form (middle), and total number of parameters.

Parameter	Form	N
f_*	DPL in M_h (equation 3) with	
	PL in z for $f_{*,0}$, M_p, γ_{lo}, γ_{hi}	8
$f_{*,\,min}$	PL in z	2
$f_{*,\,occ}$	Equation (4)	
	PL in z for M_c, μ	4
f_X	Constant	1
T_{min}	Constant	1
f_R	Constant	1
$d\log T/d\log t$	Equation (6)	3

heating occurred ∼100–300 Myr earlier in time than simple UVLF-based schemes suggest. If we are to explain the EDGES signal with the 'normal' galaxy population, we must amplify the total amount of star formation occurring at $z \gtrsim 10$, the amount of Ly-α and X-ray emission *per unit* star formation at $z > 10$, or both, relative to the baseline M17 model (similar to other UVLF-based models; Mason et al. 2015; Mashian et al. 2016; Sun & Furlanetto 2016). This is curious because, based on UVLFs alone, most theoretical models of galaxy evolution actually overproduce UV-bright galaxies at $z \sim 10$ (Oesch et al. 2018).

This problem is shown graphically in Fig. 1. Driven by the need to match the amplitude and frequency of the EDGES signal,[7] we show two primary means by which to increase the SFRD at high z: introducing a floor in the SFE, which amplifies star formation in faint galaxies only (top), and invoking a rise in the normalization of the SFE with redshift (bottom). We show the global signal results in the leftmost column, with the consequences for UVLFs in the centre and right columns, compared to a highly incomplete set of results from the recent literature[8] (van der Burg, Hildebrandt & Erben 2010; Weisz et al. 2014; Bouwens et al. 2015b; Finkelstein et al. 2015; Parsa et al. 2016; Oesch et al. 2018). Note that the Weisz et al. (2014) measurements are unlike the rest, in that they are the *reconstructed* UVLFs of local dwarfs, i.e. projections of their star formation histories back in time. In each panel, the dashed curves denote the predictions of the reference M17 model, the solid black curves indicate models that roughly match the EDGES detection, while the dotted curves merely provide another example to indicate sensitivity to the parameters.

Starting in the bottom row, we see that the $f_* \propto (1+z)^2$ dependence required to match the peak of the EDGES signal (solid curves) quickly generates tension with observed UVLFs within 500 Myr of the end of reionization, both at $z \sim 4$ and $z \sim 10$ (having anchored the model to $z \sim 6$ UVLFs). Because we have assumed monotonic redshift evolution, this model overpredicts the abundance of galaxies at $z \sim 10$ and underestimates the abundance of galaxies at $z \sim 4$. A model with $f_* \propto (1+z)$, as is expected from some analytic feedback models (Furlanetto et al. 2017), still overproduces $z \sim 10$ galaxies and underproduces $z \sim 4$ galaxies, while remaining in tension with the EDGES measurement by ∼10 MHz.

In the top row, we see that introducing a floor in the SFE can reduce tension between the UVLFs and EDGES 'for free,' i.e. without violating any observational constraints. In fact, the reconstructed star formation histories of local dwarfs seem to suggest a steepening in the faint end of the UVLF at these redshifts (Weisz et al. 2014), as shown in the top-middle panel. The dotted curve, which imposes a floor in the SFE at 5 per cent (independent of halo mass) matches the Weisz et al. (2014) points, but generates a trough in the global 21-cm slightly too early, and departs from the $z \sim 10$ UVLF fairly substantially.

Because the highest redshift UVLFs contain only the brightest galaxies, one solution to the problem at hand is to simultaneously reduce the SFE of high-mass haloes with z while amplifying the SFE of low-mass haloes. This implies a change in both the shape and the normalization of the SFE with time, which, as we will see shortly, tends towards a solution in which galaxy SFRs scale linearly with mass, i.e. the SFE becomes independent of mass. We will postpone a physical interpretation of this result to Section 4.

3.2 Fitting the EDGES signal

Due to the complex interplay between factors discussed in the previous subsection, we perform a multidimensional fit (using EMCEE; Foreman-Mackey et al. 2013) to explore a broad range of possibilities. We fit the Bouwens et al. (2015b) UVLFs from $4 \lesssim z \lesssim 8$, and the $z \sim 10$ UVLF presented in Oesch et al. (2018). Use of other UVLFs can systematically shift the inferred SFE, though the effect is minor (see e.g. Mason et al. 2015). For the global 21-cm measurement, we adopt the recovered signal shown in Figure 1 of Bowman et al. (2018), with peak amplitude $T_{21} = -530$ mK, central frequency $\nu_0 = 78.1$ MHz, width of 18.7 MHz, and flattening factor $\tau = 7$. We adopt a frequency-independent uncertainty of 100 mK (conservative root-mean squared residual after removing the foreground) across the EDGES band, deferring a more detailed treatment of covariances between model parameters, instrument, and foreground to future work.

We vary the parameters that control the SFE, including the normalization (defined as f_* at $M_h = 10^{10}\,M_\odot$), peak mass, low- and high-mass slope, and faint-end modifications (floor or steep decline), allowing each quantity to evolve with redshift. Furthermore, we allow T_{min} and f_X to vary in the fit. This results in 16 total parameters, the majority of which (14/16) are in place to describe the UVLF from $4 \lesssim z \lesssim 10$. For the excess cooling model, we also vary α, β, and z_0 (see Section 2.3.1), and for the radio background model we add f_R to the list of free parameters (see Section 2.3.2). These parameters are summarized in Table 1.

We do not vary the stellar metallicity,[9] as it is completely degenerate with the SFE in the context of our model (see section 3.4 in M17), nor do we vary the escape fraction, as the IGM must be nearly neutral at the redshifts of the EDGES signal, so the new detection offers no direct information on that parameter. We will comment on the implications for the reionization history in Section 4.

[7] None of our models match the EDGES measurement's shape in detail. We will revisit this in Section 4.

[8] We have made no effort to homogenize these data, some of which correspond to slightly different redshifts (at the level of $\Delta z \sim 0.1$) and rest-wavelengths ($\Delta\lambda \sim 100$–200 Å).

[9] For simplicity, we adopt solar metallicity ($Z = 0.02$), which means our inferred star formation rates can be scaled by the ratio of 1600 Å luminosities of stellar populations of different metallicities. However, this is only a ∼10 per cent effect (between $Z = 0.001$ and $Z = 0.02$) in the BPASS v1.0 models (Eldridge & Stanway 2009), for example, so this hardly affects our results.

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

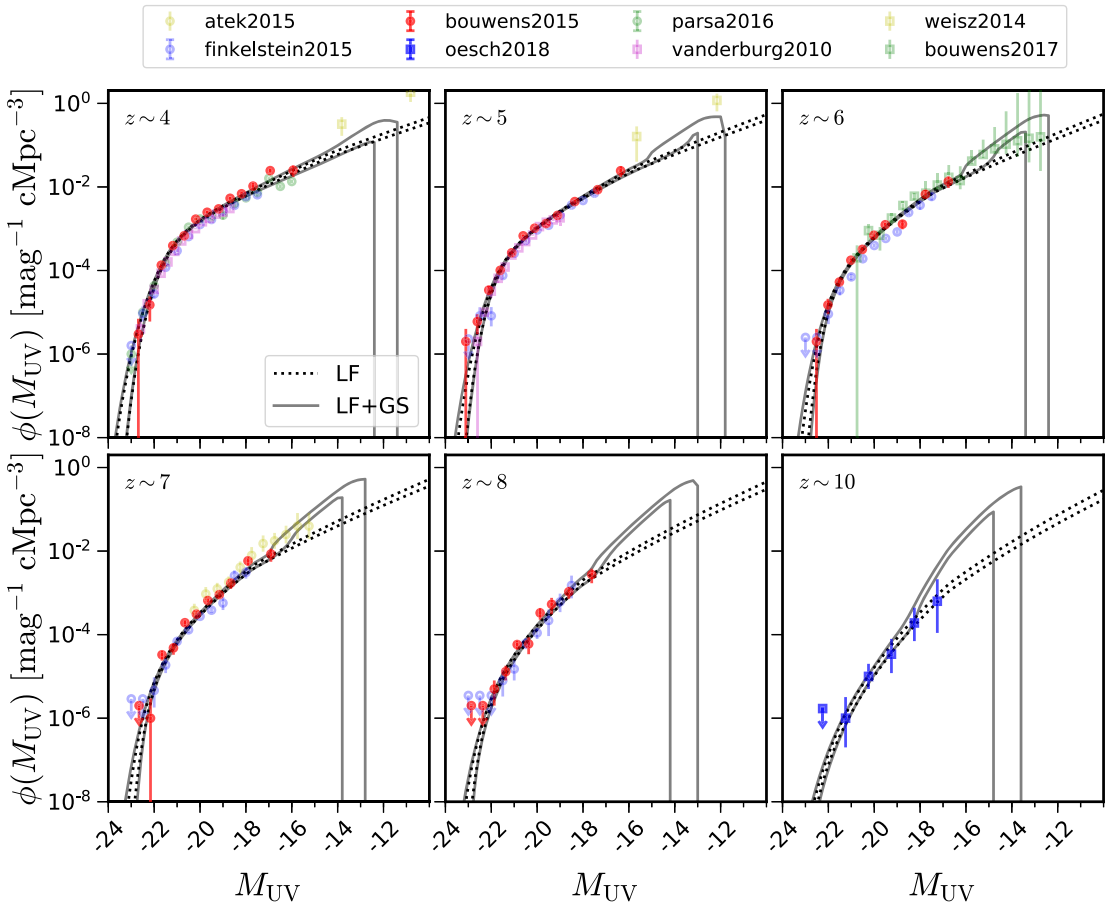

Figure 2. Recovered galaxy luminosity functions, with (solid) and without (dotted) knowledge of the global 21-cm signal. The EDGES signal demands much stronger early star formation than anticipated from an extrapolation of the observed UVLFs, which manifests in a substantial steepening of the LF in the (so far unobserved) faint regime. The pair of lines for each model bracket the 68 per cent confidence intervals in each fit. Because we have no mechanism in place to suppress galaxy formation in low-mass haloes (e.g. reionization feedback), the enhanced star formation in low-mass objects required to fit the EDGES measurement persists to arbitrarily low redshifts. Note that for simplicity, only the filled plot symbols have been used in the fit (i.e. Bouwens et al. 2015b; Oesch et al. 2018). We have also added points from Atek et al. (2015) and Bouwens et al. (2017a), which represent some of the deepest limits yet, made possible by the Hubble Frontier Fields program (Lotz et al. 2017).

3.3 Implications for galaxy formation scenarios

In Fig. 2, we show a compilation of UVLF measurements from the literature along with the results of two fits.

First, we present our reference model fits to UVLFs drawn from Bouwens et al. (2015b) and Oesch et al. (2018) at $4 \lesssim z \lesssim 10$ as dotted lines, neglecting the EDGES measurement entirely. It is an update of the M17 reference model in that it corrects UVLFs for dust attenuation (see Section 2.2) and simultaneously fits data over a broad range of redshifts.[10] We also allow the normalization of the SFE to evolve in time as a power law, as is expected in some theoretical models (e.g. Furlanetto et al. 2017). The uncertainties in ϕ are $\lesssim 0.1$ dex at fixed UV magnitude, and grow to only $\sim$0.2 dex at $z \gtrsim 15$ and $M_{\rm UV} \gtrsim -15$.

Second, in the solid lines we show the results when we simultaneously fit UVLFs and the EDGES measurement, assuming equation (6) for the cooling rate of the high-z IGM. Inclusion of the EDGES measurement requires a boost in the SFE in the faintest galaxies, as evidenced by the steepening faint-end slope. Such faint objects have thus far only been found in lensing fields, in which their abundance remains controversial (Bouwens et al. 2017a; Livermore, Finkelstein & Lotz 2017). However, the star formation histories of local dwarfs suggest that such luminosities may not be unreasonable (Weisz et al. 2014). These slopes correspond to a flat SFE, and will thus roughly mirror the slope of the DM halo mass function. Note that fitting the EDGES measurement of course does not imply such faint-end behaviour is required at $z \sim 4$. This is a byproduct of only allowing the SFE to evolve monotonically in time – in reality, reionization feedback could stifle star formation in low-mass objects at late times, even if such objects indeed harbour unexpectedly efficient star formation at $z \sim 18$.

Results for the global signal, thermal history, and radio background are shown in Fig. 3. Each row adopts a different mechanism for increasing the amplitude of the signal, including a parametric cooling excess (top), and a radio background sourced by galaxies at high z (bottom).

Focusing first on the top row, we see that our reconstructed global 21-cm signal, while broadly consistent with the EDGES measurement, cannot match its shape in detail (left-hand panel), as it lacks a flattened peak and has broader wings than the EDGES signal. The

[10]The dust correction results in a higher inferred value for the SFE in the brightest objects, at the factor of $\lesssim$2 level.

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

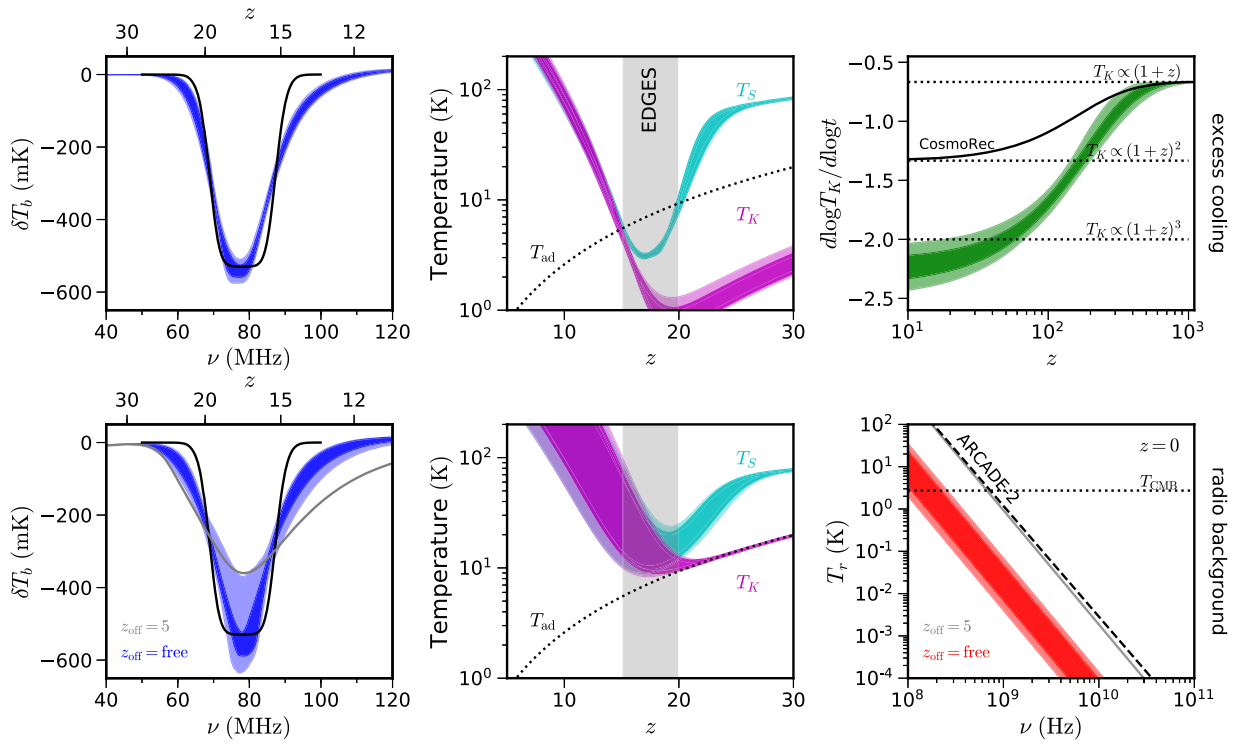

Figure 3. Reconstructed global 21-cm signal, thermal history, and radio background. *Top:* Results for an excess cooling model, in which the IGM cools following equation (6). From left to right, we show the 68 per cent and 95 per cent confidence interval for reconstructed global 21-cm signal, thermal history, and cooling rate evolution. For reference, we also show the adiabatic limit (dotted black; middle panel), which corresponds to $d\log T_{\rm K}/d\log t = -4/3$ [or $T_{\rm K} \propto (1+z)^2$; right-hand panel]. *Bottom:* Results for a radio background model, in which $L_{\rm R} \propto f_{\rm R}\dot{M}_*$. Again, we show the global signal and thermal history in the leftmost panels, but now show the $z = 0$ radio background compared to the reported ARCADE-2 excess (dashed diagonal line; Fixsen et al. 2011) and the CMB (dotted horizontal line). The grey lines indicate the best-fitting solution obtained if one allows the radio emission to continue until the calculation terminates at $z_{\rm off} = 5$, while filled contours show the results when $z_{\rm off}$ is left as a free parameter. Our best fit yields $z_{\rm off} \sim 15$ and $f_{\rm R} \sim 300$, indicating that a strong – but quickly truncated – background is required both to match the shape of the EDGES signal and also remain well below the $z = 0$ excess (right-hand panel).

temperature evolution implied by this realization is explored in the two right-hand panels. First, though the spin temperature at $z \sim 18$ is ∼3–4 K, the kinetic temperature is only ∼1–2 K, as radiative coupling is not fully complete. In order to achieve such temperatures – assuming our parametric thermal history is reasonable – cooling rates must exceed the adiabatic rate at redshifts $z \gtrsim 100$ (right-hand panel). Typically (i.e. in the absence of exotic mechanisms) cooling does not become fully adiabatic until $z \lesssim 20$.

In the bottom row of Fig. 3, we present results obtained assuming the L_R–SFR relation of equation (7). Our best-fitting normalization requires $f_{\rm R} \gtrsim 10^3$ (see Fig. 4), which is a strong requirement of star-forming galaxies. A larger value of $f_{\rm R}$ would help deepen the global 21-cm signal but would violate the ARCADE-2 excess (see right-hand panel). As a result, we also allow the boosted radio emission to terminate at some critical redshift, $z_{\rm off}$. Though an ad hoc modification of the model, interestingly, it alleviates tension at $z = 0$ while simultaneously producing a sharper global signal, as shown in the blue contours in the bottom left-hand panel of Fig. 3. Without such truncation, the radio background grows monotonically with time as the CMB decays, resulting in a signal with a broad tail to high frequencies. The thermal history in these models is consistent with standard models, implying IGM temperatures of ∼10–20 K at $z \sim 18$.

The thermal histories we recover are sensitive to the SFRD and $f_{\rm X}$. In both models, we recover $f_{\rm X} \gtrsim 10$, as shown in Fig. 4. In fact, these values are strong lower limits, given that we have assumed

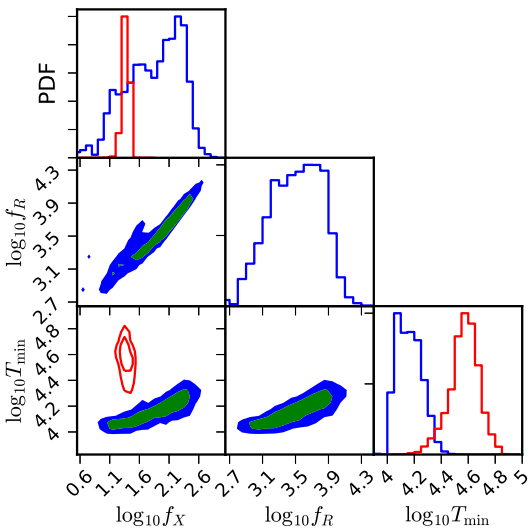

Figure 4. Constraints on f_X, f_R, and $T_{\rm min}$. Lower left three panels show joint 2D posterior distribution of the parameters, while diagonal elements show marginalized 1D constraints. The results for our phenomenological cooling model, for which $f_{\rm R}$ is unused, are shown in the red lines of the upper left and lower left-hand panels. The blue histograms and filled contours in lower left-hand panel correspond to the radio excess model with $z_{\rm off}$ left as a free parameter.

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

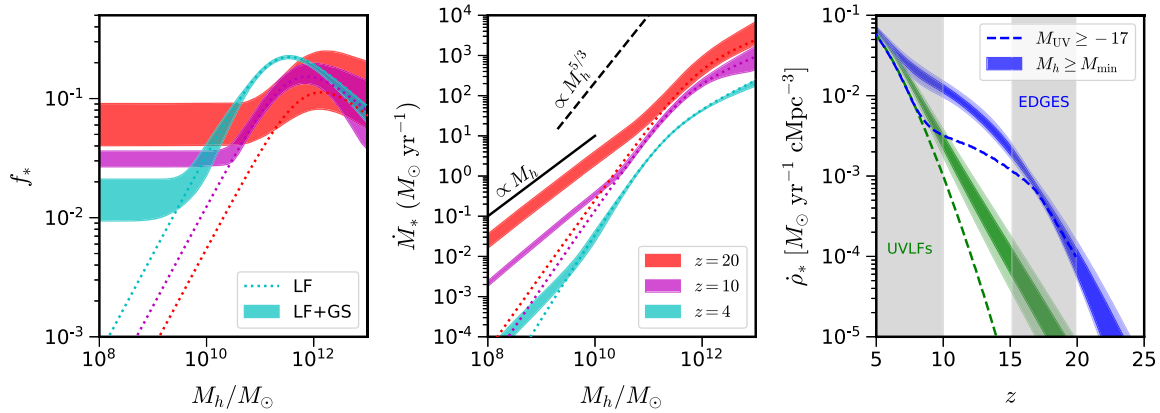

Figure 5. Recovered relationships between DM haloes and star formation in galaxies. The dotted curves correspond to UVLF-only solutions in each panel, while solid (or filled) contours indicate the results obtained when simultaneously fitting UVLFs and *EDGES*. *Left:* Efficiency of star formation, defined as the fraction of inflowing gas converted into stars. *Middle:* Relationship between SFR and halo mass. A flat SFE translates into a linear curve in this space, since DM growth rates scale roughly at $\dot{M}_{\rm h} \propto M_{\rm h}$. *Right:* Reconstructed SFRD using UVLFs only in the fit (green) versus UVLFs and the EDGES signal (blue). The dashed curves indicate the best-fitting SFRD integrated only above $M_{\rm UV} \geq -17$ for each case, while the filled contours show 68 per cent and 95 per cent confidence intervals for the total SFRD, i.e. integrating down to the minimum mass threshold.

solar metallicity (our SFRD is an overestimate if high-z galaxies are subsolar), and an unabsorbed X-ray spectrum (neutral absorption can significantly reduce heating at fixed $f_{\rm X}$; Mirocha 2014; Das et al. 2017). Should rapid heating continue throughout reionization, we should expect a strong emission feature in the global 21-cm signal at frequencies $\nu \gtrsim 100$ MHz. So far, there are only lower limits on the duration of reionization ($\Delta z \gtrsim 1$;assuming a saturated 21-cm signal) from the EDGES high-band receiver (Bowman & Rogers 2010; Monsalve et al. 2017).

Rather than presenting a large 'corner plot' illustrating the constraints on all the parameters of our model, in Fig. 5, we focus instead on the reconstructed relationships between dark matter haloes and star formation with time. These relationships give rise to the changes in the UVLF and global signal shown in Figs 2 and 3.

Starting in the left-hand panel, we show the SFE at three redshifts, comparing the full solution obtained upon fitting UVLFs and the EDGES signal (bands), as well as the standard approach with UVLFs only (dotted lines). The decline in the normalization of the SFE with time is driven by the UVLF alone (mostly the $z \sim 8$ and $z \sim 10$ points), and scales roughly as $f_* \propto (1+z)^{-1}$ in $M_{\rm h} = 10^{10}$ M$_\odot$ haloes. Though the fit could in principle have generated a model in which the slope of the SFE in low-mass objects gradually evolved in time, the actual recovered curves show a preferred departure from the power law below $M_{\rm h} \sim 10^{10}$ M$_\odot$. This is the source of the steepening in UVLFs shown in Fig. 2. A mass of 10^{10} M$_\odot$ corresponds roughly to current sensitivity limits, which, coupled with the flat SFE we infer, means the optimal fit to the UVLF and EDGES data is one that maximizes the amount of star formation occurring in small objects beyond current detection limits.

Next, in the middle panel of Fig. 5, we show the SFR–$M_{\rm h}$ relationship. Because the SFE is ultimately degenerate with our model for halo mass growth rates, it is useful to simply look at the product $\dot{M}_* = f_* \dot{M}_{\rm h}$, which is insensitive to changes in the halo growth model. In this space, a flat SFE translates to $\dot{M}_* \propto M_{\rm h}$, whereas the UVLF-only models show $\dot{M}_* \propto M_{\rm h}^{5/3}$ (roughly) to arbitrarily low masses, consistent with simple feedback arguments.

Finally, in the right-hand panel of Fig. 5, we show the reconstructed SFRD (blue contours), which shows a level of star formation at $z \sim 10$ that is ∼10 times higher than UVLF-based inferences (green contours), even when those UVLFs are extrapolated to the atomic threshold.[11] The integrated SFRD is ∼20–30x higher than the SFRD inferred from bright objects ($M_{\rm UV} \lesssim -17$) only (dashed curves). However, we emphasize that such elevated values of the SFRD need not continue to redshifts $z \lesssim 15$ – we simply have no mechanism in our model capable of stifling star formation differentially as a function of time.

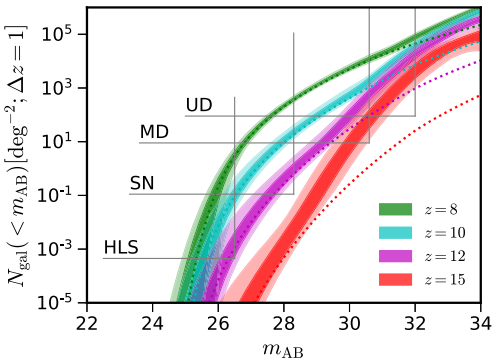

Figure 6. Predictions for galaxy counts at high z. We show the cumulative surface density of galaxies *brighter* than observed AB magnitude $m_{\rm AB}$ per unit redshift. The dotted curves show predictions based on the extrapolation of UVLFs, while the filled contours denote results obtained using UVLFs and the EDGES measurement. Four survey strategies are also shown (described in the text), with vertical lines indicating sensitivity limits, and horizontal lines highlighting the limit at which a single galaxy is found in the search area. A UD survey with *JWST* that detects ∼10 or fewer galaxies at $z \sim 12$ would be strong evidence *against* our model, or that even fainter sources are responsible for the EDGES signal.

We have intentionally limited our model to the realm of atomic cooling haloes, since they should be the most readily detectable objects in future deep surveys. In Fig. 6, we show our predictions in

[11] This more gradual decline in the SFRD is similar to the extrapolations of Madau & Dickinson (2014), and the corresponding decline of the UV luminosity density are reassuringly consistent with a 78 MHz global 21-cm signal (Madau 2018). The cause of such a smooth decline, however, is less clear, as smoothly extrapolating the SFE leads to much more rapid evolution in the SFRD (as shown in Fig. 5).

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

the context of several survey strategies, including ultra-deep (UD) and medium-deep (MD) surveys with JWST, and wide-field (WF) surveys with WFIRST, similar to their proposed supernova survey (SN) and high-latitude survey (HLS). We adopt the survey depths and areas quoted in Mason et al. (2015, their section 3.3), with the exception of the SN survey, which is like their *JWST* WF survey but covering 5x the area.

Our reference model, constructed without knowledge of the EDGES measurement, is shown in the dotted lines, and show a relatively pessimistic outlook for planned surveys at $z \gtrsim 10$–15. Such a model predicts that none of these surveys will find galaxies at $z \sim 15$, though 10–100 galaxies could be detected in each of the UD, MD, and SN surveys at $z \sim 12$.

The joint UVLF + EDGES results are shown with filled contours. If the faint end of the UVLF truly steepens, a UD survey with JWST could see ∼100 galaxies at $z \sim 12$, and perhaps ∼10 at $z \sim 15$. Unfortunately, however, the scenario in which no $z \sim 15$ galaxies are detected is still consistent with our model at the 2σ level. As a result, the strongest evidence of a steepening UVLF at high redshift would be the detection of more than ∼10 galaxies at $z \sim 12$ in a JWST UD field. In other words, strong upper limits at $N_{\rm gal} < 10$ at $z \gtrsim 12$ would point towards even fainter sources as the driving force behind the EDGES signal. In this case, 21-cm observations (with, e.g. HERA, SKA; Koopmans et al. 2015; DeBoer et al. 2017) may be the only way to constrain the properties of these sources in any detail.

4 DISCUSSION

4.1 Implications

As shown in the previous section, reconciliation of the EDGES measurement and high-z UVLF constraints is not so easily attained. We find that evolution of the SFE both in its normalization and its shape are required, if restricted to changes in the properties of atomic cooling haloes and 'normal' stellar populations.

The redshift evolution in the SFE we recover is not necessarily predicted from theoretical models. For example, while simple feedback arguments predict $f_* \propto M_{\rm h}^{2/3}$, as is roughly observed, they also predict that the normalization (at fixed mass) should *increase* with z as $f_* \propto (1+z)^{1/2}$ or $f_* \propto (1+z)$ (e.g. Dayal et al. 2014; Furlanetto et al. 2017). Physically, this rise comes from the increased binding energy of haloes (at fixed mass) at higher redshifts, enhancing the resiliency of galaxies to supernova feedback. We recover precisely the opposite trend, roughly $f_* \propto (1+z)^{-1}$. This could be an indication that our model overestimates mass accretion rates at the highest redshifts, forcing the SFE to compensate. Indeed, our model *must* overestimate inflow rates, as we have completely neglected mergers as a means by which galaxies grow. However, simulations suggest that mergers are subdominant at high z (e.g. Goerdt et al. 2015), i.e. unlikely to account for the factor of $(1+z)$ or more that is required here. Despite the decline in SFE with redshift, SFRs (at fixed $M_{\rm h}$) still rise with redshift (see Fig. 5).

Perhaps the more surprising result is the need for a change in the shape of the SFE. In our fitting, the most extreme behaviour that is allowed to occur is $f_* =$ constant with respect to $M_{\rm h}$. More gradual evolution in the slope of an unbroken power-law was a possible solution, as we allowed the slope of the SFE curve to evolve in time in our fits. Because the EDGES trough only persists to $z \sim 15$, its demand for elevated star formation ends below that point. In principle, the mechanisms that increase star formation in small galaxies could therefore also end at $z \sim 15$. We have chosen not to complicate our model by introducing an explicit parametrization like this. However, we note that even if this efficient star formation persists until $z \sim 6$, our models remain consistent with the Planck Collaboration et al. (2016) constraints on the CMB optical depth. For an escape fraction $f_{\rm esc} = 0.1$, our models generate $\tau_{\rm e} \sim 0.08$ including the faint-end UVLF steepening, and $\tau_{\rm e} \sim 0.06$ in a model with an unbroken power-law in the SFE. As a result, $f_{\rm esc} \lesssim 0.05$ is viable in models with enhanced star formation in haloes $M_{\rm h} \lesssim 10^{10}$ M$_\odot$.

This need for substantially enhanced star formation in low-mass objects could of course be ameliorated by minihaloes, which are exceedingly abundant in the early Universe. We recently found that Population III stars and their remnants (the presumed inhabitants of minihaloes) have only a subtle impact on the global signal in all but the most extreme cases (Mirocha et al. 2018). As a result, if number counts of galaxies at $z \sim 12$–15 remain low (see Fig. 6), it might imply rather extreme Pop III star formation. In this case, high-z SNs might also help distinguish the sources dominating high-z star formation (Mebane et al. 2018).

It is also possible that the somewhat contrived evolution in the SFE we require to simultaneously fit UVLFs and the EDGES signal is signifying the emergence of new source populations, rather than revealing some unexpected evolution in the properties of star-forming galaxies. Boylan-Kolchin (2018) recently pointed out that the UVLFs of globular clusters forming at high z could be comparable to those expected of galaxies themselves. This possibility is supported observationally as well, given that many objects in lensing fields are remarkably small (Bouwens et al. 2017b). Though the faint-end slope of a globular cluster population is shallower than we require (at $\alpha \sim -1.7$; Boylan-Kolchin 2018), lensing campaigns are biased towards their detection, and as a result, may obscure the true shape of the UVLF at high z (Zick, Weisz & Boylan-Kolchin 2018). Moving forward, it will be vital to understand such biases well enough to distinguish intrinsically steep UVLFs from those that have been inflated through such effects.

New source populations may also be warranted if the radio background model is to remain a viable explanation of the signal's amplitude. We have found that if star-forming galaxies are to generate this background, they must produce low-frequency emissions ∼10^3 more efficiently than galaxies today. To make matters worse, this epoch of enhanced radio emission must rapidly come to an end in order to match the shape of the EDGES signal and to fall below the $z = 0$ excess reported by ARCADE-2 (Fixsen et al. 2011). Accreting supermassive black holes may be a more viable source of this background (Ewall-Wice et al. 2018), but, as in the case of radio emission from star formation, it remains unclear what mechanisms would strongly suppress emissions at $z \lesssim 15$. This is to say nothing of the growing importance of inverse-Compton losses at high z, which likely render our estimates conservative (Sharma 2018) unless the radio background is not of astrophysical origin (Fraser et al. 2018; Pospelov et al. 2018).

4.2 Caveats

Finally, we discuss the possible shortcomings of our approach, and if any assumptions or approximations we have made could account for the discrepancy between UVLF-calibrated predictions for the global 21-cm signal (M17) and the observed global 21-cm signal reported in Bowman et al. (2018).

Our underlying model is similar to many others in the literature (e.g. Mason et al. 2015; Mashian et al. 2016; Sun & Furlanetto 2016), and agrees well with the form of the SFE inferred in these

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

works as well as their predictions for UVLFs at higher redshifts. The core assumption in each model is that star formation is fuelled by the inflow of pristine gas, proceeding continuously such that haloes of the same mass have identical assembly histories in each of their constituent components (i.e. total, gas, stellar, and metal masses). This model clearly cannot be correct in detail, as galaxies are known to exhibit scatter in their luminosities (at fixed mass), not to mention a slew of factors that can set galaxies of the same mass on entirely different long-term growth trajectories (i.e. not just minor excursions from an otherwise smooth history). However, generalizing our models to address such complexities will not obviously result in a systematic rise in star formation globally at $>$ 10, and thus may not reduce the need for efficient star formation in objects beyond current detection limits.

For example, scatter in the SFR (and thus luminosity) in haloes of fixed mass is more likely to bias the SFE to high values, if in fact the scatter is lognormal. Furthermore, because galaxies spend some of their time with SFRs less than the expected rate (given $M_{\rm h}$), it is difficult to dramatically change the SFRD. As a result, our neglect of scatter results in a model with a conservative estimate of the SFRD.

One economical way to reduce the need for enhanced star formation in objects beyond current detection limits is to invoke obscuration in the objects we do see. Though we perform a standard dust correction to correct for reddening, we have neglected the possibility that some fraction of emission in UV-detected galaxies could be completely extincted (e.g. Bowler et al. 2018). The right-hand panel of Fig. 4 shows that the required excess SFRD at $z \sim 18$ is more than an order of magnitude beyond what is possible with galaxies of similar mass to those observed at $z \sim 8$–10, which would in any case require very extreme dust corrections to completely resolve. Our models also broadly agree with constraints on the stellar-mass functions at high z (Stefanon et al. 2017), which are less susceptible to dust effects, so any bias in our inferred SFE is likely small.

The semi-empirical nature of our model permits models that could reasonably be considered contrived. However, the idea that feedback ceases to regulate star formation in sufficiently small haloes – one interpretation of a flat SFE – is not entirely unappealing. For example, the dynamical time-scale approaches the lifetimes of massive stars in haloes $M_{\rm h} \lesssim 10^9$ M$_\odot$ (Faucher-Giguère 2018). This, coupled with rapid inflow rates, could allow galaxies to form stars out of new gas before the previous generation of stars has a chance to drive proto-stellar material away via supernovae blast-waves. Alternatively, persistent contrivances may point to a more fundamental failure of models based on abundance matching and smooth inflow-driven star formation. 21-cm observations could thus provide constraints on philosophically different approaches to galaxy evolution modelling (e.g. Kelson, Benson & Abramson 2016).

Finally, our results demonstrate the importance of multiwavelength measurements of the high-z Universe for rigorously constraining the properties of luminous sources during the cosmic dawn. Ironically, early predictions for the 21-cm global signal (e.g. Furlanetto 2006; Pritchard & Loeb 2010; Mesinger et al. 2013; Mirabel et al. 2011; Fialkov et al. 2014; Mirocha, Harker & Burns 2015) yielded absorption troughs at $\lesssim$ 80 MHz in their fiducial treatments. These models, however, were not tied to observations of the galaxy population, simply assuming a single mean SFE across all haloes so could not be taken as any more than qualitative predictions. The dramatic improvements in measurements of the galaxy LF at $z > 6$ over the past several years (e.g. van der Burg et al. 2010; McLure et al. 2013; Bouwens et al. 2015b; Finkelstein et al. 2015; McLeod, McLure & Dunlop 2016; Bouwens et al. 2017a; Livermore et al. 2017; Oesch et al. 2018) required a recalibration of our expectations: like nearby galaxy populations, they require a declining SFE in small haloes – precisely those that are abundant at even higher redshifts – and this understanding allowed us to make *quantitative* predictions for their effects on the 21-cm background.

Thus it is only by *combining* UVLF measurements with the EDGES signal that we can understand the puzzle the latter poses for galaxy formation physics. Models that accurately describe the physics driving both measurements are essential for extracting the most useful and compelling implications, and improvements to the underlying theoretical models and their predictive range (including allowances for sources we have ignored in the present treatment, such as Population III star formation) will be necessary as galaxy surveys, the 21-cm data, and other measurements of the cosmic dawn improve over the next several years.

5 CONCLUSIONS

We summarize our main findings as follows:

(i) A 78 MHz feature in the global 21-cm signal is not expected based on the extrapolation of UVLFs, as outlined in M17. In order to induce the EDGES feature without violating UVLFs at $z \lesssim 10$, we must appeal to enhanced star formation in objects currently beyond detection limits. This will remain the case even in the event that the amplitude of the signal is revised downward in future studies, provided that its central frequency remains intact.

(ii) We find that, if the efficiency of star formation tends to a constant $\sim$ few per cent in haloes $M_{\rm h} \lesssim 10^{10}$ M$_\odot$, galaxies in atomic cooling haloes alone can provide enough star formation to explain the timing of the EDGES measurement. This implies a corresponding steepening in the UVLF at high z relative to past predictions, which could be tested by forthcoming observations. Detection of fewer than $\sim$10 galaxies at $z \sim 12$ in a JWST UD field would be compelling evidence that even fainter sources are required to explain the EDGES measurement, in which case 21-cm observations with interferometers like HERA, LOFAR, and the SKA may be the only way to constrain the properties of the first luminous sources in any detail.

(iii) We find that $f_{\rm X} \gtrsim 10$ is also required to fit the EDGES signal. Given that we have assumed an unabsorbed X-ray spectrum, the true value could be even higher.

(iv) For the first sources to generate a radio background capable of amplifying the global 21-cm signal, they must produce radiation at $\sim$1–2 GHz some $\sim 10^3$ times more efficiently (per unit SFR) than star-forming galaxies today. Furthermore, this efficient radio emission must terminate beyond $z \sim 15$ in order to match the shape of the signal measured by EDGES and to fall below the $z = 0$ excess.

(v) The excess cooling model only removes the need for substantial radio emission from high-z sources – the implications for the UVLF remain unchanged, as the timing of the EDGES signal, and its sharpness, provide a vital constraint on early galaxies.

ACKNOWLEDGEMENTS

The authors thank Rick Mebane, Raul Monsalve, Charlotte Mason, Aaron Ewall-Wice, and Louis Abramson for many useful discussions. This work was supported by the National Science Foundation through award AST-1636646 and by NASA through award NNX15AK80G. In addition, this work was directly sup-

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

ported by the NASA Solar System Exploration Research Virtual Institute cooperative agreement number 80ARC017M0006. We acknowledge support from a NASA contract supporting the 'WFIRST Extragalactic Potential Observations (EXPO) Science Investigation Team' (15-WFIRST15-0004), administered by GSFC. This work used computational and storage services associated with the HOFFMAN2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group, and relied on the Python packages NUMPY[12] (Oliphant 2007) and MATPLOTLIB[13] (Hunter 2007).

REFERENCES

Atek H. et al., 2015, ApJ, 800, 18
Barkana R., 2018, Nature, 555, 71
Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
Barkana R., Loeb A., 2005, ApJ, 626, 1
Behroozi P. S., Silk J., 2015, ApJ, 799, 32
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 762, L31
Bolgar F., Eames E., Hottier C., Semelin B., 2018, MNRAS, 478, 5564
Bouwens R. J. et al., 2014, ApJ, 793, 115
Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015a, ApJ, 811, 140
Bouwens R. J. et al., 2015b, ApJ, 803, 34
Bouwens R. J. et al., 2017a, ApJ, 843, 129
Bouwens R. J., Illingworth G. D., Oesch P. A., Maseda M., Ribeiro B., Stefanon M., Lam D., 2017b, preprint(arXiv:astro-ph/1711.02090)
Bowler R. A. A., Bourne N., Dunlop J. S., McLure R. J., McLeod D. J., 2018, MNRAS, 481, 1631
Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67
Bowman J. D., Rogers A. E. E., 2010, Nature, 468, 796
Boylan-Kolchin M., 2018, MNRAS, 479, 332
Brorby M., Kaaret P., Prestwich A., Mirabel I. F., 2016, MNRAS, 457, 4081
Burns J. O. et al., 2017, ApJ, 844, 33
Capak P. L. et al., 2015, Nature, 522, 455
Chen X., Miralda-Escudé J., 2004, ApJ, 602, 1
Chluba J., Thomas R. M., 2011, MNRAS, 412, 748
Chuzhoy L., Alvarez M. A., Shapiro P. R., 2006, ApJ, 648, L1
Condon J. J., Cotton W. D., Broderick J. J., 2002, AJ, 124, 675
Das A., Mesinger A., Pallottini A., Ferrara A., Wise J. H., 2017, MNRAS, 469, 1166
Dayal P., Ferrara A., Dunlop J. S., Pacucci F., 2014, MNRAS, 445, 2545
DeBoer D. R. et al., 2017, PASP, 129, 45001
Dillon J. S. et al., 2014, Phys. Rev., 89, 23002
Eldridge J. J., Stanway E. R., 2009, MNRAS, 400, 1019
Ewall-Wice A., Dillon J. S., Mesinger A., Hewitt J ., 2014, MNRAS, 441, 2476
Ewall-Wice A., Chang T.-C., Lazio J., Doré O., Seiffert M., Monsalve R. A., 2018, ApJ, 868, 63
Faucher-Giguère C.-A., 2018, MNRAS, 473, 3717
Feng C., Holder G., 2018, ApJ, 858, L17
Fialkov A., Barkana R., Pinhas A., Visbal E., 2014, MNRAS, 437, L36
Fialkov A., Barkana R., Visbal E., 2014, Nature, 506, 197
Fialkov A., Barkana R., Cohen A., 2018, Phys. Rev. Lett., 121, 011101
Field G. B., 1958, Proc. IRE, 46, 240
Finkelstein S. L. et al., 2015, ApJ, 810, 71
Fixsen D. J. et al., 2011, ApJ, 734, 5
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J ., 2013, PASP, 125, 306
Fraser S. et al., 2018, Phys. Lett., 785, 159
Furlanetto S. R., Mirocha J., Mebane R. H., Sun G., 2017, MNRAS, 472, 1576
Furlanetto S. R., 2006, MNRAS, 371, 867
Furlanetto S. R., Pritchard J. R., 2006, MNRAS, 372, 1093
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Gilfanov M., Grimm H. J., Sunyaev R., 2004, MNRAS, 347, L57
Goerdt T., Ceverino D., Dekel A., Teyssier R., 2015, MNRAS, 454, 637
Gürkan G. et al., 2018, MNRAS, 475, 3010
Haardt F., Madau P., 1996, ApJ, 461, 20
Heesen V., Brinks E., Leroy A. K., Heald G., Braun R., Bigiel F., Beck R., 2014, AJ, 147, 103
Hirata C. M., 2006, MNRAS, 367, 259
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Imara N., Loeb A., Johnson B. D., Conroy C., Behroozi P., 2018, ApJ, 854, 36
Kelson D. D., Benson A. J., Abramson L. E., 2016, preprint(arXiv:astro-ph/1610.06566)
Koopmans L. et al., 2015, Proceedings of Advancing Astrophysics with the Square Kilometer Array
Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, ApJ, 835, 113
Lotz J. M. et al., 2017, ApJ, 837, 97
Madau P., 2018, MNRAS, 480, L43
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Madau P., Fragos T., 2017, ApJ, 840, 39
Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429
Mashian N., Oesch P. A., Loeb A., 2016, MNRAS, 455, 2101
Mason C. A., Trenti M., Treu T., 2015, ApJ, 813, 21
McBride J., Fakhouri O., Ma C.-P., 2009, MNRAS, 398, 1858
McLeod D. J., McLure R. J., Dunlop J. S., 2016, MNRAS, 459, 3812
McLure R. J. et al., 2013, MNRAS, 432, 2696
Mebane R. H., Mirocha J., Furlanetto S. R., 2018, MNRAS, 479, 4544
Mesinger A., Ferrara A., Spiegel D. S., 2013, MNRAS, 431, 621
Mesinger A., Greig B., Sobacchi E., 2016, MNRAS, 459, 2342
Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64
Mineo S., Gilfanov M., Sunyaev R., 2012, MNRAS, 419, 2095
Mirabel I. F., Dijkstra M., Laurent P., Loeb A., Pritchard J. R., 2011, A&A, 528, A149
Mirocha J., Mebane R. H., Furlanetto S. R., Singal K., Trinh D., 2018, MNRAS, 478, 5591
Mirocha J., 2014, MNRAS, 443, 1211
Mirocha J., Harker G. J. A., Burns J. O., 2015, ApJ, 813, 11
Mirocha J., Furlanetto S. R., Sun G., 2017, MNRAS, 464, 1365
Mitsuda K. et al., 1984, PASJ, 36, 741
Monsalve R. A., Rogers A. E. E., Bowman J. D., Mozdzen T. J., 2017, ApJ, 847, 64
Muñoz J. B., Loeb A., 2018, Nature, 557, 684
Murray S. G., Power C., Robotham A. S. G., 2013, Astron. Comput., 3, 23
Narayanan D., Davé R., Johnson B. D., Thompson R., Conroy C., Geach J., 2018, MNRAS, 474, 1718
O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, ApJ, 807, L12
Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, ApJ, 855, 105
Oh S. P., 2001, ApJ, 553, 499
Oliphant T. E., 2007, Comput. Sci. Eng., 9, 10
Parsa S., Dunlop J. S., McLure R. J., Mortlock A., 2016, MNRAS, 456, 3194
Parsons A. R. et al., 2014, ApJ, 788, 106
Patil A. H. et al., 2017, ApJ, 838, 65
Planck Collaboration, 2016, A&A, 594, A13
Pober J. C. et al., 2015, ApJ, 809, 62
Pospelov M., Pradler J., Ruderman J. T., Urbano A., 2018, Phys. Rev. Lett., 121, 31103
Pritchard J. R., Furlanetto S. R., 2006, MNRAS, 367, 1057
Pritchard J. R., Furlanetto S. R., 2007, MNRAS, 376, 1680
Pritchard J. R., Loeb A., 2010, Phys. Rev., 82, 23006
Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901
Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19
Sharma P., 2018, MNRAS, 481, L6

[12] http://www.numpy.org/
[13] http://www.matplotlib.org/

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022

Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Singal J. et al., 2018, PASP, 130, 36001
Singh S. et al., 2017, ApJ, 845, L12
Stefanon M., Bouwens R. J., Labbé I., Muzzin A., Marchesini D., Oesch P., Gonzalez V., 2017, ApJ, 843, 36
Sun G., Furlanetto S. R., 2016, MNRAS, 460, 417
Tanaka T. L., O'Leary R. M., Perna R., 2016, MNRAS, 455, 2619
van der Burg R. F. J., Hildebrandt H., Erben T., 2010, A&A, 523, A74
Weisz D. R., Johnson B. D., Conroy C., 2014, ApJ, 794, L3
Wouthuysen S. A., 1952, AJ, 57, 31
Zick T. O., Weisz D. R., Boylan-Kolchin M., 2018, MNRAS, 477, 480
Zygelman B., 2005, ApJ, 622, 1356

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/483/2/1980/5222679 by guest on 08 February 2022