The Astrophysical Journal, 777:118 (10pp), 2013 November 10 doi:10.1088/0004-637X/777/2/118

INTERPRETING THE GLOBAL 21 cm SIGNAL FROM HIGH REDSHIFTS. I. MODEL-INDEPENDENT CONSTRAINTS

Jordan Mirocha[1,2], Geraint J. A. Harker[1,2], and Jack O. Burns[1,2]
[1] Center for Astrophysics and Space Astronomy, University of Colorado, Campus Box 389, Boulder, CO 80309, USA; jordan.mirocha@colorado.edu
[2] The NASA Lunar Science Institute, NASA Ames Research Center, Moffett Field, CA 94035, USA
Received 2013 August 5; accepted 2013 September 8; published 2013 October 21

ABSTRACT

The sky-averaged (global) 21 cm signal is a powerful probe of the intergalactic medium (IGM) prior to the completion of reionization. However, so far it has been unclear whether it will provide more than crude estimates of when the universe's first stars and black holes formed, even in the best case scenario in which the signal is accurately extracted from the foregrounds. In contrast to previous work, which has focused on predicting the 21 cm signatures of the first luminous objects, we investigate an arbitrary realization of the signal and attempt to translate its features to the physical properties of the IGM. Within a simplified global framework, the 21 cm signal yields quantitative constraints on the Lyα background intensity, net heat deposition, ionized fraction, and their time derivatives without invoking models for the astrophysical sources themselves. The 21 cm absorption signal is most easily interpreted, setting strong limits on the heating rate density of the universe with a measurement of its redshift alone, independent of the ionization history or details of the Lyα background evolution. In a companion paper, we extend these results, focusing on the confidence with which one can infer source emissivities from IGM properties.

Key words: dark ages, reionization, first stars – diffuse radiation

Online-only material: color figures

1. INTRODUCTION

Nearly all of our knowledge about the early universe comes from the observable signatures of two phase transitions: the cosmic microwave background (CMB), a byproduct of cosmological recombination at $z \sim 1100$ (Spergel et al. 2003; Komatsu et al. 2011), and Gunn–Peterson troughs in the spectra of high-z quasars (Gunn & Peterson 1965), a sign that cosmological reionization is complete by $z \gtrsim 6$. The intervening ∼Gyr, in which the first stars, black holes, and galaxies form, is very poorly understood.

Observations with the *Hubble Space Telescope* (*HST*) have begun to directly constrain galaxies well into the epoch of reionization (EoR) at redshifts possibly as high as $z \sim 10$ (e.g., Oesch et al. 2010, 2012; Bouwens et al. 2011; Zheng et al. 2012; Coe et al. 2013; Ellis et al. 2013), and upcoming facilities such as the *James Webb Space Telescope* (*JWST*) promise to extend this view even further, likely to $z \gtrsim 10$–15 (e.g., Johnson et al. 2009; Zackrisson et al. 2012). However, directly observing luminous sources at high z is not equivalent to constraining their impact on the intergalactic medium (IGM; Pritchard & Furlanetto 2007), whether it is in the form of ionization, heating, or more subtle radiative processes (e.g., the Wouthuysen–Field effect). The most promising probe of the IGM in the pre-reionization epoch is the redshifted 21 cm signal from neutral hydrogen. Its evolution over cosmic time encodes the history of heating, ionization, and Lyα emission, which in principle means that it is a probe of the background intensity at photon energies ranging from the Lyα resonance to hard X-rays (for a recent review, see Furlanetto et al. 2006a).

At stake in the quest to observe the universe in its infancy is an understanding of galaxy formation, which currently rests upon a theoretically reasonable but virtually unconstrained foundation. The first stars are expected to be very massive ($M \gtrsim 100\ M_\odot$; e.g., Haiman et al. 1996; Tegmark et al. 1997; Bromm et al. 1999; Abel et al. 2002), resulting in surface temperatures of the order of 10^5 K (Tumlinson & Shull 2000; Bromm et al. 2001; Schaerer 2002), though evidence for such objects is currently limited to abundance patterns in low-mass stars in the Milky Way (e.g., Salvadori et al. 2007; Rollinde et al. 2009). Whether or not such massive stars ever form is a vital piece of the galaxy formation puzzle, as their brief existence is expected to dramatically alter the physical conditions for subsequent star formation: first, through an intense soft-UV radiation field, which photo-ionizes (dissociates) atomic (molecular) hydrogen, and presumably via metal enrichment and thermal feedback following a supernovae explosion (see review by Bromm et al. 2009).

Even if the first stars are ∼100 $M_\odot$ and leave behind remnant black holes of comparable mass, it is difficult to reconcile the existence of $z \gtrsim 6$–7 quasars (Fan 2006; Mortlock et al. 2011), whose luminosities imply accretion onto super-massive black holes (SMBHs) with masses $M_\bullet \gtrsim 10^9\ M_\odot$, with models of growth via Eddington-limited accretion. The difficulty of growing SMBHs from modest seeds has inspired direct-collapse models (Begelman et al. 2006, 2008), which predict the formation of BHs with $M_\bullet \gtrsim 10^3\ M_\odot$ in massive, atomic-cooling dark matter halos via dynamical instabilities. These models alleviate the requirement of continual Eddington-limited accretion throughout the reionization epoch, but remain unconstrained.

JWST may be able to detect clusters of Population III (PopIII) stars at $2 \lesssim z \lesssim 7$ (Johnson 2010), PopIII galaxies and quasi-stars at $z \sim 10$–15 (Zackrisson et al. 2011; Johnson et al. 2012), and PopIII supernovae at $z \sim 15$–20 (Whalen et al. 2013a, 2013b), depending on their masses, emission properties, etc. However, the prospects for constraining the *first* generations of stars and black holes via direct detection, which likely form at higher redshifts, are bleak. The prospects for constraining the first stars and black holes *indirectly*, however, are encouraging at low radio frequencies, regardless of their detailed properties.

While the long term goal is to map the 21 cm fluctuations from the ground (a task on the horizon at $z \lesssim 10$; e.g., the Low Frequency Array (LOFAR), the Murchison Widefield Array (MWA), the Precision Array for Probing the Epoch of Reionization (PAPER), the Giant Metrewave Telescope (GMRT), and the Square Kilometer Array (SKA); Harker et al. 2010; van Haarlem et al. 2013; Bowman et al. 2013; Parsons

et al. 2010; Paciga et al. 2013; Carilli et al. 2004; Mellema et al. 2013) or space (e.g., the Lunar Radio Array (LRA); Jester & Falcke 2009) using large interferometers, in the near term, the entire $10 \lesssim z \lesssim 40$ window is likely to be accessible only to all-sky 21 cm experiments. Several challenges remain, however, from both observational and theoretical perspectives. The Earth is a sub-optimal platform for observations at the relevant frequencies ($\nu \lesssim 200$ MHz) due to radio-frequency interference and ionospheric variability (Vedantham et al. 2013), making the lunar farside a particularly appealing destination for future observatories (e.g., the LRA, the *Dark Ages Radio Explorer*; Burns et al. 2012). Some foregrounds cannot be escaped even from the lunar farside (e.g., synchrotron emission from our own galaxy) and must be removed in post-processing using sophisticated fitting algorithms (e.g., Harker et al. 2012; Liu et al. 2013). To date, ground-based 21 cm efforts have largely focused on the end of the EoR ($100 \lesssim [\nu/\mathrm{MHz}] \lesssim 200$), including lower limits on the duration of reionization (via the single-element EDGES instrument; Bowman & Rogers 2010), and constraints on the thermal and ionization history with single dish telescopes and multi-element interferometers (e.g., Paciga et al. 2013; Parsons et al. 2013). Extending this view to "cosmic dawn" requires observations below 100 MHz, a frequency range most easily explored from the radio-quiet, ionosphere-free,[3] lunar farside.

Even if the astrophysical signal is perfectly extracted from the foregrounds, it is not clear that one could glean more than gross estimates of the timing of first star and black hole formation. Although simply knowing the redshift at which the first stars and black holes form would be an enormous achievement, it is ultimately their properties that are of interest. Were the universe's first stars very massive? Did all SMBHs in the local universe form via direct collapse at high z? Could the global 21 cm signal alone rule out models for the formation of the first stars and black holes? What if independent measurements from *JWST* and/or other facilities were available?

Motivated by such questions, we turn our attention to the final stage of any 21 cm pipeline: interpreting the measurement. Rather than formulating astrophysical models and studying 21 cm realizations that result, we focus on an *arbitrary* realization of the signal and attempt to recover the properties of the universe in which it was observed. We defer a detailed discussion of how these properties of the universe (e.g., the temperature, ionized fraction, etc.) relate to astrophysical sources to Paper II (J. Mirocha et al., in preparation).

The outline of this paper is as follows. In Section 2, we introduce the physical processes that give rise to the 21 cm signal. In Section 3, we step through the three expected astrophysical features of the signal, focusing on how observational measures translate to physical properties of the universe. A discussion and conclusions are presented in Sections 4 and 5, respectively.

We adopt a cosmology with $\Omega_{\mathrm{m},0} = 0.272$, $\Omega_{\mathrm{b},0} = 0.044$, $\Omega_{\Lambda,0} = 0.728$, and $H_0 = 70.2$ km s^{-1} Mpc^{-3} throughout.

2. FORMALISM

2.1. Magnitude of the 21 cm Signal

The 21 cm transition results from hyperfine splitting in the 1S ground state of the hydrogen atom when the magnetic moments of the proton and electron flip between aligned (triplet state) and anti-aligned (singlet state). The H I brightness temperature depends sensitively on the "spin temperature," T_{S}, a 21 cm specific excitation temperature which characterizes the number of hydrogen atoms in the triplet and singlet states, $(n_1/n_0) = (g_1/g_0)\exp(-T_\star/T_{\mathrm{S}})$, where g_1 and g_0 are the degeneracies of the triplet and singlet hyperfine states, respectively, and $T_\star = 0.068$ K is the temperature corresponding to the energy difference between hyperfine levels.

The redshift evolution of the 21 cm signal, $\delta T_b(z)$, as measured relative to the CMB, also depends on the mean hydrogen ionized fraction, $\overline{x}_i$, and in general on the baryon over density and proper motions along the line of sight, though the last two effects should be negligible for studies of the all-sky spectrum, leaving (e.g., Furlanetto et al. 2006a),

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

where h is the Hubble parameter today in units of 100 km s^{-1} Mpc^{-1}, and $\Omega_{\mathrm{b},0}$ and $\Omega_{\mathrm{m},0}$ are the fractional contributions of baryons and matter to the critical energy density, respectively.

Whether the signal is seen in emission or absorption against the CMB depends entirely on the spin temperature, which is determined by the strength of collisional coupling and presence of background radiation fields,

$$T_S^{-1} \approx \frac{T_\gamma^{-1} + x_c T_K^{-1} + x_\alpha T_\alpha^{-1}}{1 + x_c + x_\alpha}, \quad (2)$$

where $T_\gamma = T_{\gamma,0}(1+z)$ is the CMB temperature, T_{K} is the kinetic temperature, and $T_\alpha \approx T_{\mathrm{K}}$ is the UV color temperature.

In general, the collisional coupling is a sum over collision-partners,

$$x_c = \sum_i \frac{n_i \kappa_{10}^i}{A_{10}}\frac{T_*}{T_\gamma}, \quad (3)$$

where n_i is the number density of species i, and $\kappa_{10}^i = \kappa_{10}^i(T_{\mathrm{K}})$ is the rate coefficient for spin de-excitation via collisions with species i. In a neutral gas, collisional coupling is dominated by hydrogen-hydrogen collisions (Allison & Dalgarno 1969; Zygelman 2005; Sigurdson & Furlanetto 2006), though hydrogen–electron collisions can become important as the ionized fraction and temperature grow (Furlanetto & Furlanetto 2007). We neglect collisional coupling due to all other species.[4]

The remaining coupling coefficient, x_α, characterizes the strength of Wouthuysen–Field coupling (Wouthuysen 1952; Field 1958),

$$x_\alpha = \frac{S_\alpha}{1+z}\frac{\widehat{J}_\alpha}{\overline{J}_\alpha}, \quad (4)$$

[3] The Moon is not truly devoid of an ionosphere—its atmosphere is characterized as a surface-bounded exosphere, whose constituents are primarily metal ions liberated by interactions with energetic particles and radiation from the Sun (e.g., Stern 1999). However, it is tenuous enough to be neglected at frequencies $\nu \gtrsim 1$ MHz.

[4] Furlanetto & Furlanetto (2007) investigated the effects of hydrogen–proton collisions on T_{S} and found that they could account for up to ~2% of the collisional coupling at $z \approx 20$, and would dominate the coupling at $z \approx 10$ in the absence of heat sources. However, an early Lyα background is expected to couple $T_{\mathrm{S}} \rightarrow T_{\mathrm{K}}$ prior to $z = 20$, and heating is expected prior $z = 10$, so protons are generally neglected in 21 cm calculations. Collisions with neutral helium atoms in the triplet state could also induce spin-exchange (Hirata & Sigurdson 2007), though the cold high-z IGM lacks the energy required to excite atoms to the triplet state. We also neglect hydrogen-deuterium collisions, whose rarity prevents any real effect on T_{S}, even though $\kappa_{10}^{\mathrm{HD}} > \kappa_{10}^{\mathrm{HH}}$ at low temperatures (Sigurdson & Furlanetto 2006). Lastly, we neglect velocity-dependent effects (Hirata & Sigurdson 2007), which introduces an uncertainty of up to a few % in the mean signal.

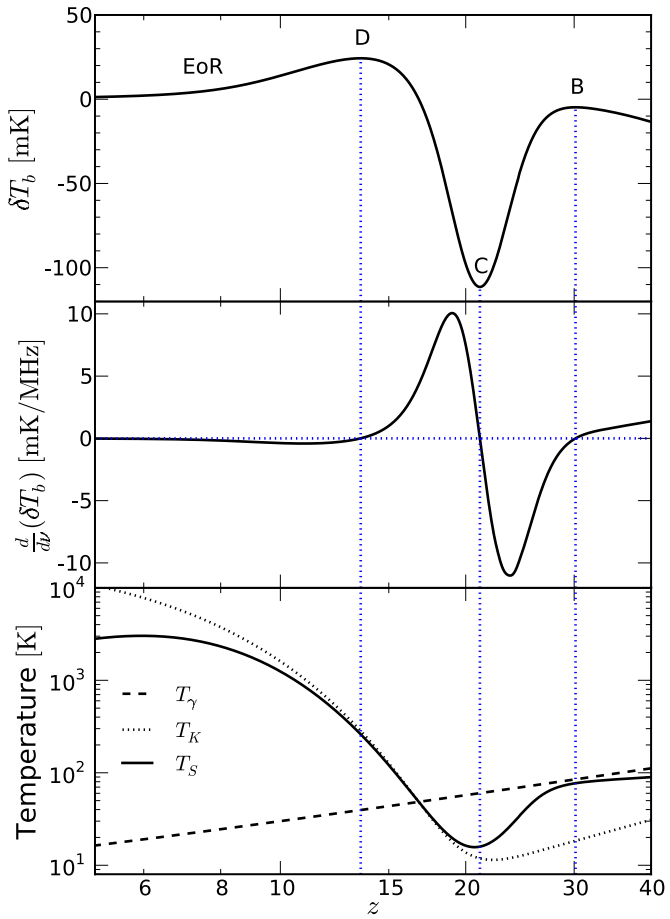

Figure 1. Example global 21 cm spectrum (top), its derivative (middle), and corresponding thermal evolution (bottom) for a model in which reionization is driven by PopII stars, and the X-ray emissivity of the universe is dominated by high-mass X-ray binaries.

(A color version of this figure is available in the online journal.)

where

$$\overline{J}_\alpha \equiv \frac{16\pi^2 T_\star e^2 f_\alpha}{27 A_{10} T_{\gamma,0} m_e c}. \quad (5)$$

$\widehat{J}_\alpha$ is the angle-averaged intensity of Lyα photons in units of s^{-1} cm^{-2} Hz^{-1} sr^{-1}, S_α is a correction factor that accounts for variations in the background intensity near line-center (Chen & Miralda-Escudé 2004; Furlanetto & Pritchard 2006; Hirata 2006), m_e and e are the electron mass and charge, respectively, f_α is the Lyα oscillator strength, and A_{10} is the Einstein A coefficient for the 21 cm transition.

2.2. Slope of the 21 cm Signal

Models for the global 21 cm signal generally result in a curve with five extrema,[5] three of which are labeled in Figure 1, roughly corresponding to the formation of the first stars (B), black holes (C), and beginning of the EoR (D). Due to the presence of strong (but spectrally smooth in principle; see Petrovic & Oh 2011) foregrounds, the "turning points" are likely the only pieces of the signal that can be reliably extracted (e.g., Pritchard & Loeb 2010; Harker et al. 2012). Our primary goal in Section 3 will be to determine the quantitative physical meaning of each feature in turn.

In preparation, we differentiate Equation (1),

$$\frac{d}{d\nu}[\delta T_b] \simeq 0.1\left(\frac{1-\overline{x}_i}{0.5}\right)\left(\frac{1+z}{10}\right)^{3/2} \times \left\{\left(\frac{T_\gamma}{T_S}\right)\left[1+\frac{3}{2}\frac{d\log T_S}{d\log t}\right] - \frac{1}{2(1-\overline{x}_i)}\left(1-\frac{T_\gamma}{T_S}\right) \times \left[1-\overline{x}_i\left(1-3\frac{d\log \overline{x}_i}{d\log t}\right)\right]\right\} \text{mK MHz}^{-1}, \quad (6)$$

making it clear that at an extremum, the following condition must be satisfied:

$$\frac{d\log T_S}{d\log t} = \frac{1}{3(1-\overline{x}_i)}\left(\frac{T_S}{T_\gamma}-1\right) \times \left[1-\overline{x}_i\left(1-3\frac{d\log \overline{x}_i}{d\log t}\right)\right] - \frac{2}{3}. \quad (7)$$

We can obtain a second independent equation for the spin-temperature rate of change by differentiating Equation (2),

$$\frac{d\log T_S}{d\log t} = \left[1+x_{\rm tot}\left(\frac{T_\gamma}{T_K}\right)\right]^{-1}\left\{\frac{x_{\rm tot}}{(1+x_{\rm tot})}\frac{d\log x_{\rm tot}}{d\log t} \times \left[1-\left(\frac{T_\gamma}{T_K}\right)\right] + x_{\rm tot}\frac{d\log T_K}{d\log t}\left(\frac{T_\gamma}{T_K}\right) - \frac{2}{3}\right\}, \quad (8)$$

where $x_{\rm tot} = x_c + x_\alpha$, such that

$$\frac{d\log x_{\rm tot}}{d\log t} = x_{\rm tot}^{-1}\left[\sum_i x_c^i \frac{d\log x_c^i}{d\log t} + x_\alpha \frac{d\log x_\alpha}{d\log t}\right]. \quad (9)$$

Expanding out the derivatives of the coupling terms, we have

$$\frac{d\log x_\alpha}{d\log t} = \frac{d\log \widehat{J}_\alpha}{d\log t} + \frac{d\log S_\alpha}{d\log T_K}\frac{d\log T_K}{d\log t} + \frac{2}{3} \quad (10)$$

and

$$\frac{d\log x_c^i}{d\log t} = \frac{d\log \kappa_{10}^i}{d\log T_K}\frac{d\log T_K}{d\log t} \pm \frac{d\log x_e}{d\log t} - \frac{4}{3}, \quad (11)$$

where the second to last term is positive for H–H collisions and negative for H–e^- collisions.

As in Furlanetto (2006) and Pritchard & Furlanetto (2007), we adopt a two-zone model in which the volume filling fraction of H II regions, x_i, is treated separately from the ionization in the bulk IGM, parameterized by x_e. The mean ionized fraction is then $\overline{x}_i = x_i + (1-x_i)x_e$. This treatment is motivated[6]

[5] We neglect the first and last features of the signal in this paper. The lowest redshift feature marks the end of reionization, and while its frequency derivative is zero, so is its amplitude, making its precise location difficult to pinpoint. The highest redshift feature is neglected because it is well understood theoretically and should occur well before the formation of the first luminous objects (though exotic physics such as dark-matter annihilation could complicate this, e.g., Furlanetto et al. 2006b).

[6] Our motivation for the logarithmic derivative convention is primarily compactness, though the non-dimensionalization of derivatives is convenient for comparing the rate at which disparate quantities evolve. For reference, the logarithmic derivative of a generic function of redshift with respect to time, $d\log w/d\log t = b$, implies $w(z) \propto (1+z)^{-3b/2}$ under the high-z approximation, $H(z) \approx H_0\Omega_{\rm m,0}^{1/2}(1+z)^{3/2}$, which is accurate to better than $\sim$0.5% for all $z > 6$. For example, the CMB cools as $d\log T_\gamma/d\log t = -2/3$.

by the fact that $\delta T_b = 0$ in H II regions, thus eliminating the need for a detailed treatment of the temperature and ionization evolution, but beyond H II regions the gas is warm and only partially ionized (at least at early times) so we must track both the kinetic temperature and electron density in order to compute the spin temperature.

3. CRITICAL POINTS IN THE 21 cm HISTORY

From the equations of Section 2, it is clear that in general turning points in the 21 cm signal probe a set of eight quantities, $\boldsymbol{\theta} = \{x_i, x_e, T_{\rm K}, \widehat{J}_\alpha, x_i', x_e', T_{\rm K}', \widehat{J}_\alpha'\}$, where primes represent logarithmic time derivatives. Given a perfect measurement of the redshift and brightness temperature, $(z, \delta T_b)$, at a turning point, the system is severely underdetermined with two equations (Equations (1) and (7)) and eight unknowns. Without independent measurements of the thermal and/or ionization history and/or Lyα background intensity, no single element of $\boldsymbol{\theta}$ can be constrained unless one or more assumptions are made to reduce the dimensionality of the problem.

The most reasonable assumptions at our disposal are as follows.

1. The volume-filling factor of the H II regions, x_i, and the ionized fraction in the bulk IGM, x_e, are both negligible, as are their time derivatives, such that $\overline{x}_i = d\log\overline{x}_i/d\log t = 0$.
2. There are no heat sources, such that the universe's temperature is governed by pure adiabatic cooling after decoupling at $z_{\rm dec} \simeq 150$ (Peebles 1993), i.e., $d\log T_{\rm K}/d\log t = -4/3$.
3. Lyα coupling is strong, i.e., $x_\alpha \gtrsim 1$, such that $T_{\rm S} \rightarrow T_{\rm K}$, and the dependencies on $\widehat{J}_\alpha$ no longer need be considered.

These assumptions are expected to be valid at $z \gtrsim 10$, $z \gtrsim 20$ and $z \lesssim 10$, respectively, according to typical models (e.g., Furlanetto 2006; Pritchard & Loeb 2010). However, since it may be impossible to verify their validity from the 21 cm signal alone, we will take care in the following sections to state explicitly how each assumption affects inferred values of $\boldsymbol{\theta}$. We will now examine each feature of the signal in turn.

3.1. Turning Point B: End of the Dark Ages

Prior to the formation of the first stars, the universe is neutral to a part in $\sim 10^4$ (e.g., `RECFAST`, `HyRec`, `CosmoRec`; Seager et al. 1999, 2000; Ali-Haïmoud & Hirata 2010; Chluba & Thomas 2011), such that a measurement of δT_b probes $T_{\rm S}$ directly via Equation (1),

$$T_{\rm S} \leqslant T_\gamma \left[1 - \frac{\delta T_b}{9\ {\rm mK}}(1+z)^{-1/2}\right]^{-1}, \quad (12)$$

where the $\leqslant$ symbol accounts for the possibility that $\overline{x}_i > 0$ (a non-zero ionized fraction always acts to reduce the amplitude of the signal). For the first generation of objects, we can safely assume $\overline{x}_i \ll 1$, and interpret a measurement of the brightness temperature as a proper constraint on $T_{\rm S}$ (rather than an upper limit). We will relax this requirement in Section 3.2.

If $T_{\rm S}$ and $T_{\rm K}$ are both known, Equation (2) yields the total coupling strength, $x_{\rm tot}$. However, the contribution from collisional coupling is known as a function of redshift for a neutral adiabatically cooling gas, and can simply be subtracted from $x_{\rm tot}$ to yield x_α, and thus $\widehat{J}_\alpha$ (via Equation (4)). The top panel of Figure 2 shows lines of constant $\log_{10}(J_\alpha/J_{21})$, where $J_\alpha = h\nu_\alpha \widehat{J}_\alpha$ and $J_{21} = 10^{-21}$ erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}, given the redshift and brightness temperature of turning point B, $\delta T_b(z_{\rm B})$. From Equations (7) and (8), we can also constrain the rate of change in the background Lyα intensity (Equation (10)), as shown in the bottom panel of Figure 2.

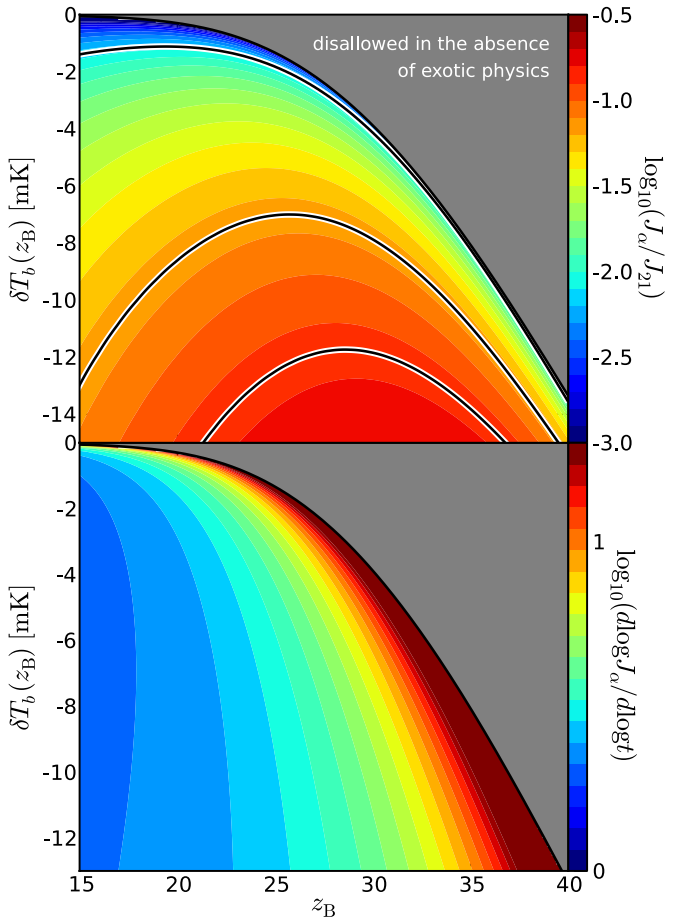

Figure 2. Values $J_\alpha = h\nu_\alpha \widehat{J}_\alpha$ and $d\log J_\alpha/d\log t$ that give rise to turning point B at position $(z_{\rm B}, \delta T_b(z_{\rm B}))$. The color scale shows the value of J_α (top panel, in units of $J_{21} = 10^{-21}$ erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}), and $d\log\widehat{J}_\alpha/d\log t$ (bottom panel) required for turning point B to appear at the corresponding position in the $(z_{\rm B}, \delta T_b(z_{\rm B}))$ plane, under the assumptions given in Section 3.1. The gray shaded region is excluded unless heating occurs in the dark ages. For reference, the highlighted black contours represent Lyα fluxes (assuming a flat spectral energy distribution at energies between Lyα and the Lyman limit, $h\nu_\alpha \leqslant h\nu \leqslant h\nu_{\rm LL}$), corresponding to Lyman–Werner band fluxes of $J_{\rm LW}/J_{21} = \{10^{-2}, 10^{-1}, 2\times10^{-1}\}$ (from top to bottom), which roughly bracket the range of fluxes expected to induce negative feedback in minihalos at $z \sim 30$ (Haiman et al. 2000).

(A color version of this figure is available in the online journal.)

In the event that heating has already begun (rendering $T_{\rm K}(z)$ unknown), interpreting turning point B becomes more complicated.[7] Now, x_α will be overestimated, given that a larger (unknown) fraction of $x_{\rm tot}$ is due to collisional coupling. Uncertainty in $T_{\rm K}$ propagates to S_α, meaning x_α can only be considered to provide an upper limit on the product $S_\alpha\widehat{J}_\alpha$, rather than $\widehat{J}_\alpha$ alone. The interpretation of the turning point condition (Equation (7)) becomes similarly complicated if no knowledge of $T_{\rm K}(z)$ is assumed.

[7] We deem such a scenario "exotic" because it requires heat sources prior to the formation of the first stars. Heating via dark matter annihilation is one example of such a heating mechanism (Furlanetto et al. 2006b).

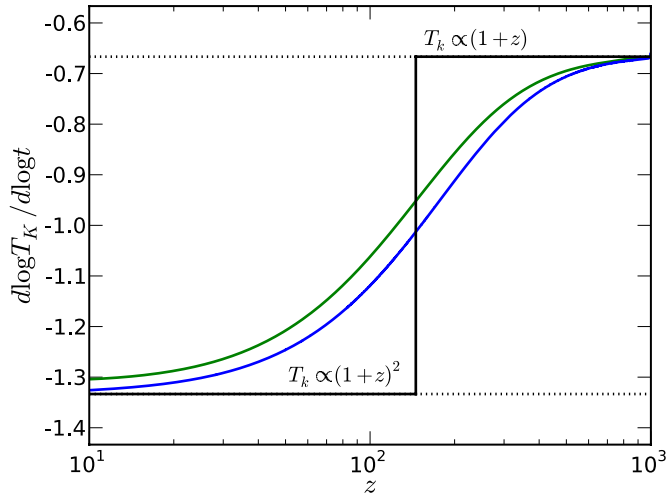

Figure 3. Cooling rate of the universe under different assumptions. The black line is an approximate analytic solution (Peebles 1993), while the blue and green lines are numerical solutions. The blue curve considers cooling via radiative recombination, collisional excitation and ionization, and the Hubble expansion, and heating via Compton scattering. The green line is an even more detailed numerical solution obtained with the `CosmoRec` code (Chluba & Thomas 2011), which includes a multi-level atom treatment and many radiative transfer effects.

(A color version of this figure is available in the online journal.)

3.2. Turning Point C: Heating Epoch

In the general case where Hubble cooling and heating from astrophysical sources must both be considered, the temperature evolution can be written as

$$\frac{d\log T_{\rm K}}{d\log t} = \frac{\tau_H}{\tau_X} - \mathcal{C}, \tag{13}$$

where we have defined a characteristic heating timescale $\tau_X^{-1} \equiv \epsilon_{\rm heat}/e_{\rm int}$, where $e_{\rm int}$ is the gas internal energy, $\epsilon_{\rm heat}$ and $\mathcal{C}$ are the heating and cooling rate densities, respectively, and $\tau_H^{-1} = 3H(z)/2$ is a Hubble time at redshift z in a matter-dominated universe.

In a neutral medium, the solution to Equation (13) for an arbitrary $\epsilon_{\rm heat}$ is

$$T_{\rm K}(z) = \mathcal{C}_1^{-1} \int_z^\infty \epsilon_{\rm heat}(z')\frac{dt}{dz'}dz' + T_{\gamma,0}\frac{(1+z)^2}{1+z_{\rm dec}}, \tag{14}$$

where $\mathcal{C}_1 \equiv 3\bar{n}_{\rm H}^0(1+y)k_{\rm B}/2$, $k_{\rm B}$ is Boltzmann's constant, $\bar{n}_{\rm H}^0$ is the hydrogen number density today, y is the primordial helium abundance (by number), and the second term represents the adiabatic cooling limit.

To move forward analytically, we again adopt the maximal cooling rate, $\mathcal{C} = 4/3$. Detailed calculations with `CosmoRec` indicate that such a cooling rate is not achieved until $z \lesssim 10$ in the absence of heat sources, which means we *overestimate* the cooling rate, and thus *underestimate* $T_{\rm K}$ at all redshifts. This lower bound on the temperature is verified in Figure 3, in which we compare three different solutions for the cooling rate density evolution.

In order for the 21 cm signal to approach emission, the temperature must be increasing relative to the CMB,[8] i.e., $\tau_H/\tau_X > 4/3$, meaning the existence of turning point C, at redshift $z_{\rm C}$, alone gives us a lower limit on $\epsilon_{\rm heat}(z_{\rm C})$. The

[8] Though see Section 3.2.2 for an alternative scenario.

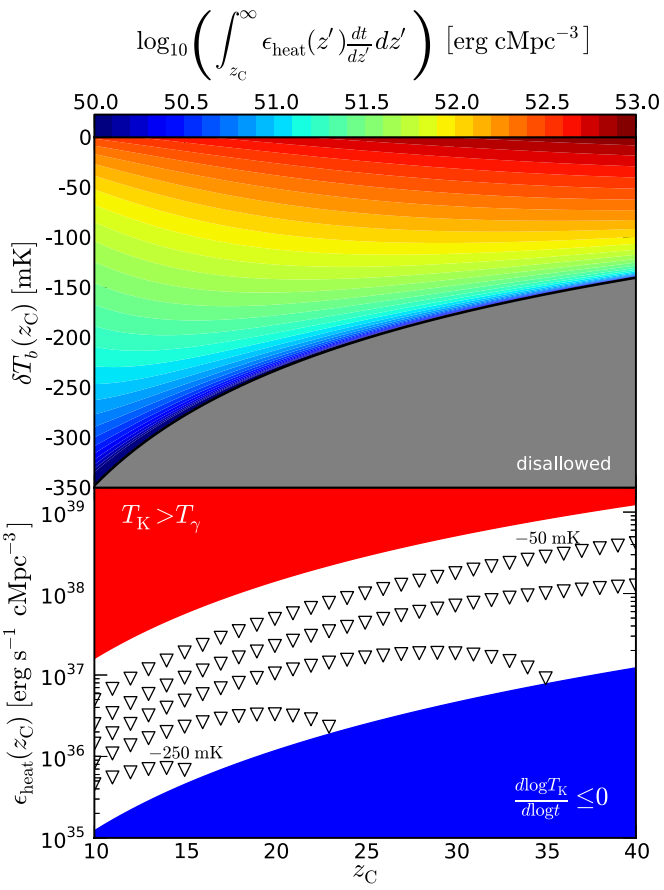

Figure 4. Top: constraints on the cumulative energy deposition as a function of the redshift and brightness temperature of turning point C. The gray region is disallowed because it requires cooling to be more rapid than Hubble (adiabatic) cooling. Bottom: constraints on the co-moving heating rate density (cMpc^{-3} means co-moving Mpc^{-3}) as a function of $z_{\rm C}$ alone. The blue region includes heating rate densities insufficient to overcome the Hubble cooling, while the red region is inconsistent with the existence of an absorption feature at $z_{\rm C}$ because such heating rates would instantaneously heat $T_{\rm K}$ above T_γ. The triangles, plotted in increments of 50 mK between $\delta T_b = \{-250, -50\}$ mK, show how a measurement of $\delta T_b(z_{\rm C})$, as opposed to $z_{\rm C}$ alone, enables more stringent upper limits on the heating rate density.

(A color version of this figure is available in the online journal.)

detection of the absorption signal (regardless of its amplitude) also requires the kinetic temperature to be cooler than the CMB temperature. If we assume a "burst" of heating, $\epsilon_{\rm heat} \rightarrow \epsilon_{\rm heat}\delta(z - z_{\rm C})$, where δ is the Dirac delta function, and require $T_{\rm K} < T_\gamma$, we can solve Equation (13) and obtain an upper limit on the co-moving heating rate density. The bottom panel of Figure 4 shows the upper and lower limits on $\epsilon_{\rm heat}$ as a function of $z_{\rm C}$ alone.

A stronger upper limit on $\epsilon_{\rm heat}(z_{\rm C})$ is within reach, however, if we can accurately measure the brightness temperature of turning point C. Given that $\delta T_b(z_{\rm C})$ provides an upper limit on $T_{\rm S}$ for all values of $\bar{x}_i$ (Equation (12)), and an absorption signal requires $T_{\rm K} < T_{\rm S} < T_\gamma$, we can solve Equation (14) assuming $T_{\rm K} < T_{\rm S}$, and once again assume a burst of heating to get a revised upper limit on $\epsilon_{\rm heat}(z_{\rm C})$.

In general, turning point C yields an upper limit (again because we have assumed $\mathcal{C} = 4/3$) on the *integral* of the heating rate density (Equation (14)), which is seen in the top

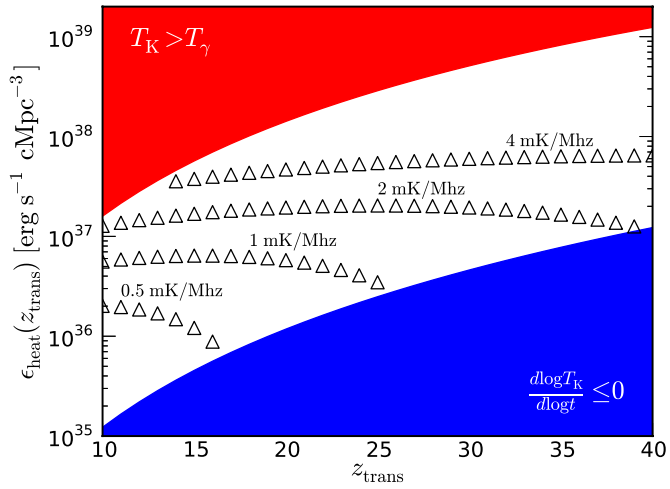

Figure 5. Constraints on the co-moving heating rate density (once again, cMpc^{-3} means co-moving Mpc^{-3}) as a function of the absorption–emission transition redshift, $z_{\rm trans}$, and the slope of the 21 cm signal at that redshift. As in Figure 4, the blue region indicates heating rates insufficient to overcome the Hubble cooling, while the red region denotes heating rates that would instantaneously heat $T_{\rm K}$ above T_γ. The triangles show how measuring the slope of the signal at $z_{\rm trans}$ can provide a lower limit on $\epsilon_{\rm heat}$.

(A color version of this figure is available in the online journal.)

panel of Figure 4.[9] This upper limit is independent of the ionization history, since any ionization reduces the amplitude of $T_{\rm S}$, thus lessening the amount of heating required to explain an absorption feature of a given depth. The only observational constraints available to date are consistent with X-ray heating of the IGM at $z \gtrsim 8$ (Parsons et al. 2013).

3.2.1. From Absorption to Emission

If heating persists, and the universe is not yet reionized, the 21 cm signal will eventually transition from absorption to emission. At this time, coupling is expected to be strong such that at the precise redshift of the transition, $z_{\rm trans}$, Equation (6) takes special form since $T_{\rm S} \simeq T_{\rm K} = T_\gamma$,

$$\frac{d}{d\nu}\left[\delta T_b\right] \simeq 0.1\left(\frac{1-x_i}{0.5}\right)\left(\frac{1+z_{\rm trans}}{10}\right)^{3/2} \times \left[1+\frac{3}{2}\frac{d\log T_{\rm K}}{d\log t}\right] \text{mK MHz}^{-1}. \quad (15)$$

That is, if we can measure the slope at the absorption–emission transition, we obtain a lower limit on the heating rate density. Our inferred heating rate density would be exact if $\overline{x}_i$ were identically zero, but for $\overline{x}_i > 0$, the slope provides a lower limit. This is illustrated in the Figure 5.

3.2.2. Could the Absorption Feature be Ionization-driven?

The absorption feature of the all-sky 21 cm signal is generally expected to occur when X-rays begin heating the IGM (e.g., Ricotti et al. 2005; Ciardi et al. 2010). However, this feature could also be produced given sufficient ionization, which similarly acts to drive the signal toward emission (albeit by reducing the absolute value of δT_b rather than increasing $T_{\rm S}$). We now assess whether or not such a scenario could produce turning point C while remaining consistent with current constraints from the Thomson optical depth to the CMB (τ_e; Dunkley et al. 2009; Larson et al. 2011; Bennett et al. 2013).

We assume that coupling is strong, $T_{\rm S} \simeq T_{\rm K}$, and that the universe cools adiabatically (i.e., the extreme case where turning point C is *entirely* due to ionization), so that a measurement of δT_b is a direct proxy for the ionization fraction (via Equation (1)). If we adopt a *tanh* model of reionization, parameterized by the midpoint of reionization, $z_{\rm rei}$, and its duration, $\Delta z_{\rm rei}$, we can solve Equation (1) at a given $\delta T_b(z_{\rm C})$ for $\overline{x}_i(z_{\rm C})$. Then, we can determine the ($z_{\rm rei}$, $\Delta z_{\rm rei}$) pair, and thus entire ionization history $\overline{x}_i(z)$, consistent with our measure of $\overline{x}_i(z_{\rm C})$. Computing the Thomson optical depth is straightforward once $\overline{x}_i(z)$ is in hand—we assume He III reionization occurs at $z = 3$, and that He II and hydrogen reionization occur simultaneously.

At a turning point, however, Equation (7) must also be satisfied. This results in a unique track through $(z, \delta T_b)$ space corresponding to values of $z_{\rm C}$ and $\delta T_b(z_{\rm C})$ that are consistent with both $\overline{x}_i(z_{\rm C})$ and its time derivative for a given *tanh* model. Figure 6 shows the joint ionization and 21 cm histories consistent with *WMAP* 9 constraints on τ_e (Bennett et al. 2013).

This technique is limited because it assumes a functional form for the ionization history that may be incorrect, in addition to the fact that we are only using two points in the fit—the first being $z_{\rm rei}$, at which point $\overline{x}_i = 0.5$ (by definition), and the second being $\overline{x}_i(z_{\rm C})$ as inferred from $\delta T_b(z_{\rm C})$. However, it does show that reasonable reionization scenarios could produce turning point C, although at later times (lower redshifts) than typical models (where turning point C is a byproduct of heating) predict.

3.3. Turning Point D: Reionization

In principle, turning point D could be due to a sudden decline in the Lyα background intensity, which would cause $T_{\rm S}$ to decouple from $T_{\rm K}$ and re-couple to the CMB. Alternatively, turning point D could occur if heating subsided enough for the universe to cool back down to the CMB temperature. However, the more plausible scenario is that coupling continues between $T_{\rm S}$ and $T_{\rm K}$, heating persists, and the signal "saturates," i.e., $1 - T_\gamma/T_{\rm S} \approx 1$, in which case the brightness temperature is a direct proxy for the volume filling factor of H II regions.[10]

If saturated, Equation (7) becomes

$$\frac{\overline{x}_i}{1-\overline{x}_i}\frac{d\log \overline{x}_i}{d\log t} \simeq \left(\frac{T_\gamma}{T_{\rm K}}\right)\frac{d\log T_{\rm K}}{d\log t} - \frac{1}{3}. \quad (16)$$

Even in the saturated regime, the first term on the right-hand side cannot be discarded since we have assumed nothing about $d\log T_{\rm K}/d\log t$.

Many authors have highlighted the 21 cm emission signal as a probe of the ionization history during the EoR (e.g., Pritchard et al. 2010; Morandi & Barkana 2012). Rather than dwell on it, we simply note that if 21 cm measurements of the EoR signal are accompanied by independent measures of $\overline{x}_i$, in principle one could glean insights into the thermal history from turning point D as well.

[9] We express our results in units of erg cMpc^{-3} to ease the conversion between $\epsilon_{\rm heat}$ and the X-ray emissivity, $\hat{\epsilon}_X$ (see Section 4.2). For reference, 10^{51} erg cMpc$^{-3} \simeq 10^{-4}$ eV baryon^{-1}.

[10] If the signal is not yet saturated, a measurement of turning point D instead yields an upper limit on $\overline{x}_i$.

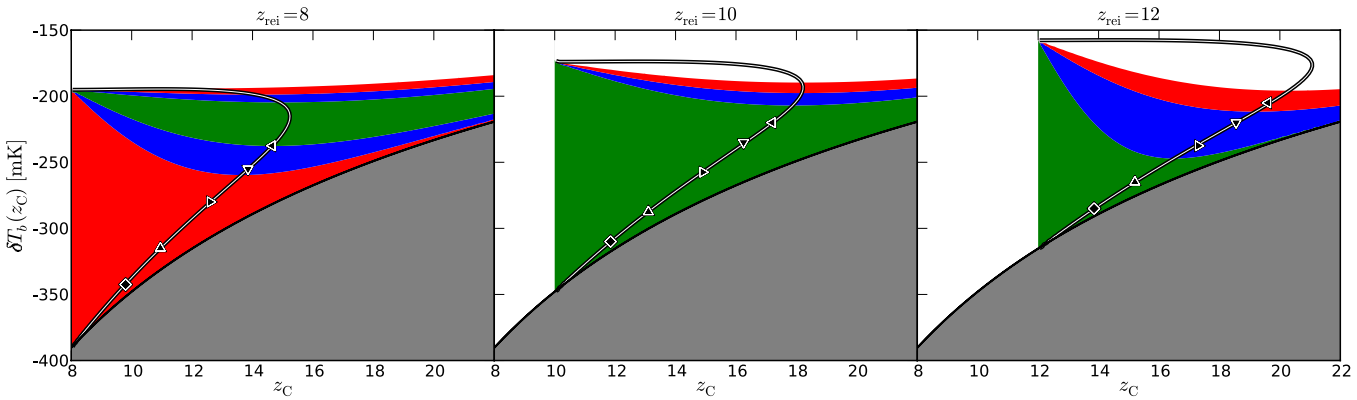

Figure 6. Plausibility of an ionization-driven absorption feature assuming *tanh* models of reionization with $z_{\rm rei} = 8$, 10, and 12 from left to right. The filled contours denote measures of $\delta T_b(z)$ (and thus $\overline{x}_i$ assuming an adiabatically cooling universe) that correspond to ionization histories consistent with *WMAP* 9 values of τ_e (Bennett et al. 2013) at the 1, 2, and 3-σ level (green, blue, and red, respectively). However, while the filled contours denote plausible reionization scenarios, not all of them would induce a turning point in the global 21 cm signal. The white contour denotes $(z, \delta T_b)$ pairs where Equation (7) is satisfied exactly, meaning $(z, \delta T_b)$ points lying within the filled contours *and* along the white contour mark locations where turning point C would be a produced by ionization and also be consistent with the CMB constraint. The symbols denote *tanh* models with $\Delta z_{\rm rei} = 1, 2, 4, 6, 8$ (diamond, upward/rightward/downward/leftward-facing triangles, respectively). The values of $\Delta z_{\rm rei} \leqslant 7.9$ are consistent with the most conservative (model-dependent) constraints from South Pole Telescope (via the kinetic Sunyaev–Zel'dovich (SZ) effect; Zahn et al. 2012), which assume no prior knowledge of angular correlations in the cosmic infrared background and thermal SZ power.

(A color version of this figure is available in the online journal.)

4. DISCUSSION

4.1. A Shift in Methodology

The redshifted 21 cm signal has been studied by numerous authors in the last 10–15 years. Efforts have concentrated on identifying probable sources of Lyα, Lyman-continuum, and X-ray photons at high z, and then solving for their combined influence on the thermal and ionization state of gas surrounding individual objects (e.g., Madau et al. 1997; Thomas & Zaroubi 2008; Chen et al. 2008; Venkatesan & Benson 2011), or the impact of populations of sources on the global properties of the IGM (e.g., Choudhury & Ferrara 2005; Furlanetto 2006; Pritchard & Loeb 2010). It has been cited as a probe of the first stars (Barkana & Loeb 2005), stellar-mass black holes and active galactic nuclei (e.g., Mirabel et al. 2011; Mcquinn 2012; Tanaka et al. 2012; Fragos et al. 2013; Mesinger et al. 2013), which primarily influence the thermal history through X-ray heating, but could contribute non-negligibly to reionization (e.g., Dijkstra et al. 2004; Pritchard et al. 2010; Morandi & Barkana 2012). More recently, more subtle effects have come into focus, such as the relative velocity difference between baryons and dark matter, which delays the formation of the first luminous objects (Tseliakhovich & Hirata 2010; McQuinn & O'Leary 2012; Fialkov et al. 2012).

Forward modeling of this sort, where the input is a set of astrophysical parameters and the output is a synthetic global 21 cm spectrum, is valuable because it (1) identifies the processes that most affect the signal, (2) has so far shown that a 21 cm signal should exist given reasonable models for early structure formation, and (3) that the signal exhibits the same qualitative features over a large subset of parameter space. However, this methodology yields no information about how unique a given model is.

We have taken the opposite approach. Rather than starting from an astrophysical model and computing the resulting 21 cm spectrum, we begin with an arbitrary signal characterized by its extrema, and identify the IGM properties that would be consistent with its observation. The advantage is that (1) we have a mathematical basis to accompany our intuition about which physical processes give rise to each feature of the signal, (2) we can see how reliably IGM properties can be constrained given a perfect measurement of the signal, and (3) we can predict which models will be degenerate without even computing a synthetic 21 cm spectrum.

4.2. An Example History

In our analysis, we have found that the 21 cm signal provides more than coarse estimates of when the first stars and black holes form. Turning points B, C, and D constrain (quantitatively) the background Lyα intensity, cumulative energy deposition, and mean ionized fraction, respectively, as well as their time derivatives, as summarized in Table 1. For concreteness, we will now revisit each feature of the signal for an assumed realization of the 21 cm spectrum, and demonstrate how each can be interpreted in terms of model-independent IGM properties.

We will assume the same realization of the signal as is shown in Figure 1, with turning points B, C, and D at $(z, \delta T_b/{\rm mK})$ of $(30.2, -4.8)$, $(21.1, -112)$, and $(13.5, 24.5)$, respectively, and absorption–emission transition at $z_{\rm trans} = 15$, $d(\delta T_b)/d\nu = 4.3$ mK MHz^{-1}. At a glance, the 21 cm realization shown in Figure 1 indicates that the universe's first stars form at $z \gtrsim 30$, the first black holes form at $z \gtrsim 21$, and that reionization has begun by $z \gtrsim 13.5$. Global feedback models such as those presented in Tanaka et al. (2012) are inconsistent with this realization of the signal, as they predict $T_{\rm K} > T_\gamma$ at $z \gtrsim 20$.

More quantitatively, from Figure 2 we have an upper limit on the Lyα background intensity of $\widehat{J}_\alpha(z_{\rm B})/J_{21} \geqslant 10^{-1.8}$ and its time rate-of-change, $d \log \widehat{J}_\alpha / d \log t \simeq 11.2$. Moving on to turning point C (Figure 4), the kinetic temperature is constrained between $9 \lesssim T_{\rm K}/{\rm K} \lesssim 16$, meaning that the cumulative energy deposition must be $\int \epsilon_{\rm heat} dt \leqslant 10^{51.9}$ erg cMpc^{-3}. In the absence of any ionization, a minimum heating rate density of $\epsilon_{\rm heat} \geqslant 10^{36.1}$ erg s^{-1} cMpc^{-3} is required to produce turning point C, and a maximum of $\epsilon_{\rm heat} \leqslant 10^{38.2}$ erg s^{-1} cMpc^{-3} is imposed given the existence of the absorption feature.

The slope of the signal as it crosses $\delta T_b = 0$ is $\delta T_b' = 4.3$ mK MHz^{-1}, corresponding to a lower limit on the heating rate density of $\epsilon_{\rm heat} \geqslant 10^{37.6}$ erg s^{-1} cMpc^{-3} (Figure 5). Finally, at turning point D, the ionized fraction must be $\overline{x}_i \leqslant 0.24$

Table 1
Features of the Global 21 cm Signal

Feature	Measurement	Assumptions	Yield	Section	Equations	Figures
B	z_B	...	Lower limit on redshift of first star formation	3.1	...	...
B	$\delta T_b(z_B)$	$\overline{x}_i = \epsilon_{heat} = 0$	$\widehat{J}_\alpha(z_B)$, $\widehat{J}'_\alpha(z_B)$	3.1	1–4, 7–11	2
C	z_C	...	Upper limit on $\epsilon_{heat}(z_C)$	3.2	14	4
C	z_C	$\overline{x}_i = 0$	Lower limit on redshift of first X-ray source formation	3.2	...	...
C	z_C	$\overline{x}_i = 0$	Lower limit on $\epsilon_{heat}(z_C)$	3.2	13	3, 4
C	$\delta T_b(z_C)$	...	Improved upper limit on $\epsilon_{heat}(z_C)$	3.2	1, 7, 13, 14	4
C	$\delta T_b(z_C)$	$\epsilon_{heat} = 0$	Rule out reionization scenario?	3.2.2	7	6
Transition	z_{trans}	$T_S = T_K$	Upper limit on $\int \epsilon_{heat} dt$	3.2.1	14	5
Transition	$\frac{d}{d\nu}[\delta T_b](z_{trans})$	$T_S = T_K$	Lower limit on $\epsilon_{heat}(z_{trans})$	3.2.1	15	5
D	z_D	...	Start of EoR	3.3	...	...
D	$\delta T_b(z_D)$	...	Upper limit on $\overline{x}_i(z_D)$	3.3	1	...
D	$\delta T_b(z_D)$	$T_S = T_K \gg T_\gamma$	$\overline{x}_i(z_D)$, joint constraint on $\overline{x}'_i(z_D)$, $T_K(z_D)$, and $T'_K(z_D)$	3.3	1, 7, 16	...

Notes. Constraints on IGM properties from critical points in the global 21 cm signal. Each block focuses on a single feature of the signal (denoted in Column 1) and from left to right reports how a given measurement (Column 2; e.g., the feature's redshift, z) under some set of assumptions (Column 3) would be interpreted (Column 4). The corresponding section of the text, as well as any equations and figures relevant to the given feature are listed in Columns 5–7, respectively. Within each block, elements appear in order of increasing complexity (in terms of the measurement difficulty and number of assumptions) from top to bottom.

(Equation (1) when $T_S \gg T_\gamma$). An ionization-driven turning point C can be ruled out by Figure 6, since the amount of ionization required to produce $(z_C, \delta T_b(z_C)) = (21.1, -112)$ leads to τ_e values inconsistent with the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) at the $>3\sigma$ level, for *tanh* models with $8 \leqslant z_{rei} \leqslant 12$.

With limits on $\widehat{J}_\alpha$, ϵ_{heat}, $\overline{x}_i$, and their derivatives, the next step is to determine how each quantity relates to astrophysical quantities. Typically, models for the global 21 cm signal relate the emissivity of the universe to the cosmic star-formation rate density (SFRD) via simple scalings of the form $\hat{\epsilon}_{i,\nu}(z) \propto f_i \dot{\rho}_*(z) I_\nu$ (e.g., Furlanetto 2006; Pritchard & Loeb 2010), in which case the parameters of interest are f_i, which converts a star formation rate into a bolometric energy output in band i (generally split between Lyα, soft-UV, and X-ray photons), the SFRD itself, $\dot{\rho}_*$, and the spectral energy distribution (SED) of luminous sources being modeled, I_ν.

Given that soft-UV photons have very short mean-free-paths in a neutral medium, a determination of $d \log \overline{x}_i / d \log t$ is likely to be an accurate tracer of the soft-UV ionizing emissivity of the universe, $\hat{\epsilon}_{ion}$. However, the same is not true of photons emitted between Ly-n resonances and hard X-ray photons, which can travel large distances before being absorbed, where they predominantly contribute to Wouthuysen–Field coupling and heating, respectively. Because of this, translating $\widehat{J}_\alpha$ and ϵ_{heat} measurements to their corresponding emissivities, $\hat{\epsilon}_\alpha$ and $\hat{\epsilon}_X$, is non-trivial. In general, the accuracy with which one can convert $\widehat{J}_\alpha$ (ϵ_{heat}) to $\hat{\epsilon}_\alpha$ ($\hat{\epsilon}_X$) depends on the redshift-evolution of the co-moving bolometric luminosity and the SED of sources, I_ν.

For a 0th order estimate, we will assume that sources have a flat spectrum between the Lyα resonance and the Lyman limit, and neglect "injected photons," i.e., those that redshift into a higher Ly-n resonance and (possibly) cascade through the Lyα resonance. If $\hat{\epsilon}_\alpha \propto N_\alpha \dot{\rho}_*$, where N_α is the number of photons emitted between $\nu_\alpha \leqslant \nu \leqslant \nu_{LL}$ per baryon, then

$$\dot{\rho}_*(z) \approx 10^{-5}\, M_\odot\ \mathrm{yr}^{-1}\ \mathrm{cMpc}^{-3} \left(\frac{9690}{N_\alpha}\right)\left(\frac{J_\alpha}{J_{21}}\right)\left(\frac{1+z}{30}\right)^{-1/2}, \quad (17)$$

where we have scaled N_α to a value appropriate for low-mass PopII stars (Barkana & Loeb 2005).

Similarly, if we assume that a fraction $f_{X,h} = 0.2$ of the X-ray emissivity is deposited as heat (appropriate for the $E \gtrsim 0.1$ keV limit in a neutral medium; Shull & van Steenberg 1985), and normalize by the local L_X–SFR relationship (e.g., Mineo et al. 2012, who found $L_{0.5-8\,\mathrm{keV}} = 2.61 \times 10^{39}$ erg s^{-1} ($M_\odot$ yr^{-1})), we have

$$\dot{\rho}_*(z) \approx 2 \times 10^{-2}\, M_\odot\ \mathrm{yr}^{-1}\ \mathrm{cMpc}^{-3}\, f_X^{-1} \times \left(\frac{0.2}{f_{X,h}}\right)\left(\frac{\epsilon_{heat}}{10^{37}\ \mathrm{erg\ s^{-1}\ cMpc^{-3}}}\right),$$

where we subsume all uncertainty in the normalization between L_X and $\dot{\rho}_*$, the SED of X-ray sources, and radiative transfer effects into the factor f_X.

If these approximate treatments are sufficient, then measures of J_α provide two-dimensional constraints on $\dot{\rho}_*$ and N_α, and measures of ϵ_{heat} constrain $\dot{\rho}_*$ and f_X.[11] However, given the long mean free paths of X-rays and photons in the $\nu_\alpha \leqslant \nu \leqslant \nu_{LL}$ band, the estimates above are likely to be inadequate. It is the primary goal of a forthcoming paper (J. Mirocha et al., in preparation) to characterize uncertainties in these estimates that arise due to two major unknowns: (1) redshift evolution in the ionizing emissivity of UV and X-ray sources, and (2) their SEDs.

4.3. Synergies with Upcoming Facilities

The prospects for synergies are most promising for turning point D, which is predicted to occur at $z \lesssim 15$, coinciding with the *JWST* window and current and upcoming campaigns to measure the 21 cm power spectrum. *JWST* will probe the high-z galaxy population even more sensitively than *HST* (e.g., Robertson et al. 2013), which may allow degeneracies between the star-formation history and other parameters to be

[11] Here we have assumed that high-mass X-ray binaries are the only source of X-rays, when in reality the heating may be induced by a variety of sources. Other candidates include X-rays from "miniquasars" (e.g., Kuhlen & Madau 2005), inverse Compton scattered CMB photons off high energy electrons accelerated in supernovae remnants (Oh 2001), or shock heating (e.g., Gnedin & Shaver 2004; Furlanetto & Loeb 2004).

broken (e.g., the f_i normalization factors). However, our focus in this paper is on model-independent quantities—the issue of degeneracy among astrophysical parameters will be discussed in Paper II.

In terms of model-independent quantities, current and upcoming facilities will benefit global 21 cm measurements by constraining the ionization history. For example, one can constrain $\overline{x}_i(z)$ via observations of Lyα-emitters (LAEs; e.g., Malhotra & Rhoads 2006; McQuinn et al. 2007; Mesinger & Furlanetto 2008), the CMB through τ_e and the kinetic Sunyaev–Zel'dovich effect (Zahn et al. 2012), or via measurements of the 21 cm power spectrum, which reliably peaks when $\overline{x}_i \simeq 0.5$ (Lidz et al. 2008). However, like the global signal, power spectrum measurements yield upper limits on $\overline{x}_i$, since they assume $T_S \gg T_\gamma$, which may not be the case. Constraints from LAEs require no such assumption, and instead set lower limits on $\overline{x}_i$, since our ability to see Lyα emission from galaxies at high-z depends on the *minimum* size of an H II region required for Lyα photons to escape. Limits on $\overline{x}_i(z)$ out to $z \sim 10$–15 would yield a prediction for the amplitude of turning point D, which, in conjunction with a global 21 cm measurement, could validate or invalidate the $T_S \gg T_\gamma$ assumption often adopted for EoR work. In addition, one could determine if ionization-driven absorption features are even remotely feasible (Section 3.2.2).

4.4. Caveats

Simple models for the global 21 cm signal rely on the assumption that the IGM is well approximated as a two-phase medium, one phase representing H II regions, and the other representing the bulk IGM. As reionization progresses, the distinction between these two phases will become tenuous, owing to a warming and increasingly ionized IGM whose properties differ little from an H II region. Even prior to reionization the global approximation may be inadequate depending on the distribution of luminous sources. If exceedingly rare sources dominate ionization and heating, we would require a more detailed treatment (a problem recently addressed in the context of helium reionization by Davies & Furlanetto 2012).

Eventually, simple models must also be calibrated by more sophisticated simulations. This has been done to some extent already in the context of 21 cm fluctuations, with good agreement so far between semi-analytic and numerical models (Zahn et al. 2011). However, analogous comparisons for the global signal have yet to be performed rigorously. The limiting factor is that a large volume must be simulated in order to avoid cosmic variance, but the spatial resolution required to simultaneously resolve the first galaxies becomes computationally restrictive.

Finally, though we included an analysis of the absorption–emission transition point, $z_{\rm trans}$, in truth, the slope measured from this feature will be correlated with the positions of the turning points. The most promising foreground removal studies rely on parameterizing the signal as a simple function (e.g., spline), meaning the slope at $z_{\rm trans}$ is completely determined by the positions of the turning points and the function used to represent the astrophysical signal.

5. CONCLUSIONS

In this paper, we have addressed one tier of the 21 cm interpretation problem: identifying the physical properties of the IGM that can be constrained uniquely from a measurement of the all-sky 21 cm signal. Our main conclusions are as follows.

1. The first feature of the global signal, turning point B, provides a lower limit on the redshift at which the universe's first stars formed. However, more quantitatively, its position in $(z, \delta T_b)$ space measures the background Lyα intensity, $\widehat{J}_\alpha$, and its time derivative, respectively, assuming a neutral, adiabatically cooling medium.
2. The absorption feature, turning point C, is most likely a probe of accretion onto compact objects considering the τ_e constraint from the CMB. As a result, it provides a lower limit on the redshift when the first X-ray emitting objects formed. Even if the magnitude of the absorption trough cannot be accurately measured, a determination of $z_{\rm C}$ alone sets strong upper and lower limits on the heating rate density of the universe, $\epsilon_{\rm heat}(z_{\rm C})$. If the absorption feature is deep ($\delta T_b(z_{\rm C}) \lesssim -200$ mK) and occurs late ($z \lesssim 15$), it could be a byproduct of reionization.
3. The final feature, turning point D, indicates the start of the EoR, and traces the mean ionized fraction of the universe and its time derivative. In general, it also depends on the spin-temperature evolution, though it is expected that at this stage the signal is fully saturated. Without independent constraints on the thermal history, $\delta T_b(z_{\rm D})$ provides an upper limit on the mean ionized fraction, $\overline{x}_i$.

In general, the relationship between IGM diagnostics (such as $\widehat{J}_\alpha$ and $\epsilon_{\rm heat}$) and the properties of the astrophysical sources themselves (like $\dot{\rho}_*$, N_α, and f_X) is expected to be complex. In a forthcoming paper, we compare simple analytic arguments (e.g., those used in Section 4.2) with the results of detailed numerical solutions to the cosmological radiative transfer equation in order to assess how accurately the global 21 cm signal can constrain the universe's luminous sources.

The authors thank the anonymous referee, whose suggestions helped improve the quality of this manuscript, and acknowledge the LUNAR consortium,[12] headquartered at the University of Colorado, which is funded by the NASA Lunar Science Institute (via Cooperative Agreement NNA09DB30A) to investigate concepts for astrophysical observatories on the Moon.

REFERENCES

Abel, T., Bryan, G. L., & Norman, M. L. 2002, Sci, 295, 93
Ali-Haïmoud, Y., & Hirata, C. M. 2010, PhRvD, 82, 063521
Allison, A. C., & Dalgarno, A. 1969, ApJ, 158, 423
Barkana, R., & Loeb, A. 2005, ApJ, 626, 1
Begelman, M. C., Rossi, E. M., & Armitage, P. J. 2008, MNRAS, 387, 1649
Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ApJ, 737, 90
Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, 31
Bowman, J. D., & Rogers, A. E. E. 2010, Natur, 468, 796
Bromm, V., Coppi, P. S., & Larson, R. B. 1999, ApJL, 527, L5
Bromm, V., Kudritzki, R. P., & Loeb, A. 2001, ApJ, 552, 464
Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F. 2009, Natur, 459, 49
Burns, J. O., Lazio, J., Bale, S., et al. 2012, AdSpR, 49, 433
Carilli, C. L., Furlanetto, S., Briggs, F., et al. 2004, NewAR, 48, 1029
Chen, X., Chen, X., Miralda-Escudé, J., & Miralda-Escudé, J. 2008, ApJ, 684, 18
Chen, X., & Miralda-Escudé, J. 2004, ApJ, 602, 1
Chluba, J., & Thomas, R. M. 2011, MNRAS, 412, 748
Choudhury, T. R., & Ferrara, A. 2005, MNRAS, 361, 577
Ciardi, B., Salvaterra, R., & Di Matteo, T. 2010, MNRAS, 401, 2635
Coe, D., Zitrin, A., Carrasco, M., et al. 2013, ApJ, 762, 32
Davies, F. B., & Furlanetto, S. R. 2012, MNRAS, arXiv:1209.4900
Dijkstra, M., Haiman, Z., & Loeb, A. 2004, ApJ, 613, 646
Dunkley, J., Komatsu, E., Nolta, M. R., et al. 2009, ApJS, 180, 306
Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, ApJL, 763, L7

[12] http://lunar.colorado.edu

Fan, X. 2006, NewAR, 50, 665
Fialkov, A., Barkana, R., Tseliakhovich, D., & Hirata, C. M. 2012, MNRAS, 424, 1335
Field, G. B. 1958, PIRE, 46, 240
Fragos, T., Lehmer, B. D., Naoz, S., Zezas, A., & Basu-Zych, A. R. 2013, ApJL, 776, L31
Furlanetto, S. R. 2006, MNRAS, 371, 867
Furlanetto, S. R., & Furlanetto, M. R. 2007, MNRAS, 374, 547
Furlanetto, S. R., & Furlanetto, M. R. 2007, MNRAS, 379, 130
Furlanetto, S. R., & Loeb, A. 2004, ApJ, 611, 642
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006a, PhR, 433, 181
Furlanetto, S. R., Oh, S. P., & Pierpaoli, E. 2006b, PhRvD, 74, 103502
Furlanetto, S. R., & Pritchard, J. R. 2006, MNRAS, 372, 1093
Gnedin, N. Y., & Shaver, P. A. 2004, ApJ, 608, 611
Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633
Haiman, Z., Abel, T., & Rees, M. J. 2000, ApJ, 534, 11
Haiman, Z., Thoul, A. A., & Loeb, A. 1996, ApJ, 464, 523
Harker, G., Zaroubi, S., Bernardi, G., et al. 2010, MNRAS, 405, 2492
Harker, G. J. A., Pritchard, J. R., Burns, J. O., & Bowman, J. D. 2012, MNRAS, 419, 1070
Hirata, C. M. 2006, MNRAS, 367, 259
Hirata, C. M., & Sigurdson, K. 2007, MNRAS, 375, 1241
Jester, S., & Falcke, H. 2009, NewAR, 53, 1
Johnson, J. L. 2010, MNRAS, 404, 1425
Johnson, J. L., Greif, T. H., Bromm, V., Klessen, R. S., & Ippolito, J. 2009, MNRAS, 399, 37
Johnson, J. L., Whalen, D. J., Fryer, C. L., & Li, H. 2012, ApJ, 750, 66
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Kuhlen, M., & Madau, P. 2005, MNRAS, 363, 1069
Larson, D., Dunkley, J., Hinshaw, G., et al. 2011, ApJS, 192, 16
Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., & Hernquist, L. 2008, ApJ, 680, 962
Liu, A., Pritchard, J. R., Tegmark, M., & Loeb, A. 2013, PhRvD, 87, 043002
Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
Malhotra, S., & Rhoads, J. E. 2006, ApJL, 647, L95
Mcquinn, M. 2012, MNRAS, 426, 1349
McQuinn, M., Hernquist, L., Zaldarriaga, M., & Dutta, S. 2007, MNRAS, 381, 75
McQuinn, M., & O'Leary, R. M. 2012, ApJ, 760, 3
Mellema, G., Koopmans, L. V. E., Abdalla, F. A., et al. 2013, ExA, 36, 235
Mesinger, A., Ferrara, A., & Spiegel, D. S. 2013, MNRAS, 431, 621
Mesinger, A., & Furlanetto, S. R. 2008, MNRAS, 386, 1990
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
Morandi, A., & Barkana, R. 2012, MNRAS, 424, 2551
Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Natur, 474, 616
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2010, ApJL, 709, L16
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012, ApJ, 759, 135
Oh, S. P. 2001, ApJ, 553, 499
Paciga, G., Albert, J. G., Bandura, K., et al. 2013, MNRAS, 433, 639
Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, AJ, 139, 1468
Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2013, arXiv:1304.4991
Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton, NJ: Princeton Univ. Press)
Petrovic, N., & Oh, S. P. 2011, MNRAS, 413, 2103
Pritchard, J. R., & Furlanetto, S. R. 2007, MNRAS, 376, 1680
Pritchard, J. R., & Loeb, A. 2010, PhRvD, 82, 23006
Pritchard, J. R., Loeb, A., & Wyithe, J. S. B. 2010, MNRAS, 408, 57
Ricotti, M., Ostriker, J. P., & Gnedin, N. Y. 2005, MNRAS, 357, 207
Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71
Rollinde, E., Vangioni, E., Maurin, D., et al. 2009, MNRAS, 398, 1782
Salvadori, S., Schneider, R., & Ferrara, A. 2007, MNRAS, 381, 647
Schaerer, D. 2002, A&A, 382, 28
Seager, S., Sasselov, D. D., & Scott, D. 1999, ApJL, 523, L1
Seager, S., Sasselov, D. D., & Scott, D. 2000, ApJS, 128, 407
Shull, J. M., & van Steenberg, M. E. 1985, ApJ, 298, 268
Sigurdson, K., & Furlanetto, S. R. 2006, PhRvL, 97, 091301
Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175
Stern, S. A. 1999, RvGeo, 37, 453
Tanaka, T., Perna, R., & Haiman, Z. 2012, MNRAS, 425, 2974
Tegmark, M., Silk, J., Rees, M. J., et al. 1997, ApJ, 474, 1
Thomas, R. M., & Zaroubi, S. 2008, MNRAS, 384, 1080
Tseliakhovich, D., & Hirata, C. 2010, PhRvD, 82, 083520
Tumlinson, J., & Shull, J. M. 2000, ApJL, 528, L65
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
Vedantham, H. K., Koopmans, L. V. E., de Bruyn, A. G., et al. 2013, arXiv:1306.2172
Venkatesan, A., & Benson, A. 2011, MNRAS, 417, 2264
Whalen, D. J., Even, W., Lovekin, C. C., et al. 2013a, ApJ, 768, 195
Whalen, D. J., Fryer, C. L., Holz, D. E., et al. 2013b, ApJL, 762, L6
Wouthuysen, S. A. 1952, AJ, 57, 31
Zackrisson, E., Rydberg, C.-E., Schaerer, D., Östlin, G., & Tuli, M. 2011, ApJ, 740, 13
Zackrisson, E., Zitrin, A., Trenti, M., et al. 2012, MNRAS, 427, 2212
Zahn, O., Mesinger, A., McQuinn, M., et al. 2011, MNRAS, 414, 727
Zahn, O., Reichardt, C. L., Shaw, L., et al. 2012, ApJ, 756, 65
Zheng, W., Postman, M., Zitrin, A., et al. 2012, Natur, 489, 406
Zygelman, B. 2005, ApJ, 622, 1356