Constraining the Tail End of Reionization Using Ly\(\alpha\) Transmission Spikes

Enrico Garaldi\textsuperscript{1,2} \textregistered, Nickolay Y. Gnedin\textsuperscript{3,4,5}, and Piero Madau\textsuperscript{6} \textcopyright

\textsuperscript{1}Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany; egaraldi@mpa-garching.mpg.de
\textsuperscript{2}Argelander Institut für Astronomie der Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany
\textsuperscript{3}Particle Astrophysics Center, Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
\textsuperscript{4}Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, IL 60637, USA
\textsuperscript{5}Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637, USA
\textsuperscript{6}Department of Astronomy & Astrophysics, University of California, 1156 High Street, Santa Cruz, CA 95064, USA

Received 2019 February 20; revised 2019 March 20; accepted 2019 March 21; published 2019 April 30

Abstract

We investigate Ly\(\alpha\) transmission spikes at \(z > 5\) in synthetic quasar spectra and discuss their connection to the properties of the intergalactic medium and their ability to constrain reionization models. We use state-of-the-art radiation-hydrodynamic simulations from the Cosmic Reionization On Computers series to predict the number of transmission spikes as a function of redshift, both in the ideal case of infinite spectral resolution and in a realistic observational setting. Transmission spikes are produced in highly ionized underdense regions located in the vicinity of UV sources. We find that most of the predicted spikes are unresolved by current observations and show that our mock spectra are consistent with observations of the quasar ULAS J1120+0641 in about 15\% of the realizations. The spike height correlates with both the gas density and the ionized fraction, but the former link is erased when synthetic spectra are smoothed to realistically achievable spectral resolutions. There exists a linear relationship between spike width and the extent of the associated underdense region, with a slope that is redshift dependent. In agreement with observations, the spike transmitted flux is suppressed at small distance from bright galaxies as these reside in overdense regions. We argue that this anticorrelation can be used to constrain large-scale density modes.

Key words: dark ages, reionization, first stars – galaxies: high-redshift – intergalactic medium – methods: numerical – quasars: absorption lines

1. Introduction

The hydrogen stored in the intergalactic medium (IGM) transformed from completely neutral to a highly ionized plasma at redshift \(5.5 \lesssim z \lesssim 10\) (Planck Collaboration et al. 2018) in what is known as the Epoch of Reionization (EoR). It is widely believed that the ionizing photons that caused this transition were produced by early star-forming galaxies (see, e.g., Madau et al. 1999; Gnedin 2000; Haardt & Madau 2012), with active galactic nuclei (Haa\textsuperscript{rdt} & Madau 1996; Kulkarni et al. 2018) and X-ray binaries (Eide et al. 2018) playing only a minor role at redshifts \(z > 5\) (but see Madau & Haardt 2015).

There are both empirical and theoretical arguments supporting this picture. Observationally, the Cosmic Microwave Background anisotropies place the mid point of reionization at \(7.64 \pm 0.74\) (if symmetric, Planck Collaboration et al. 2018), while the evolution of the effective optical depth of the Ly\(\alpha\) forest (e.g., Fan et al. 2006) and the sudden change in the density of detected Ly\(\alpha\)-emitting systems (e.g., Ota et al. 2010; Pentericci et al. 2011; Mason et al. 2018) constrain the tail end of reionization (where individual ionized bubbles overlap) to occur at redshift \(5.5 \lesssim z \lesssim 6.5\). Additional constraints at various cosmic times have been obtained from modeling the near zones of high-redshift quasars (e.g., Schroeder et al. 2013; Davies et al. 2018) and from the dark pixel statistics in quasar absorption spectra (e.g., McGreer et al. 2011).

The amount of UV photons produced by galaxies, however, appears to be insufficient to fully ionize intergalactic hydrogen if their escape fraction into the IGM is of order a few percent, similar to the one typically observed in the local universe (see, e.g., Marchi et al. 2017). Some lower-redshift galaxies exhibit escape fraction in excess of 10\% (and as large as \(~60\%,\) Vanzella et al. 2018), but it remains unclear whether these are representative of the sources driving the EoR. Models typically assume larger escape fractions at higher redshift, although a consensus on the physical mechanism causing such evolution has not yet been reached.

Historically, high-redshift quasars have been one of the most powerful probes of the EoR, as they provide bright flashlights that illuminate the IGM along the line of sight to Earth. The most prominent and ubiquitous absorption line in quasar spectra is the Ly\(\alpha\) transition of neutral hydrogen. The cross section for such process is very large, so that even a modest hydrogen neutral fraction of \(x_\text{HI} \sim 10^{-4}\) produces complete absorption. For this reason, Ly\(\alpha\) studies have been historically limited to the postreionization universe and mainly to the range \(3 \lesssim z \lesssim 4.5\), where the superposition of multiple individual absorption features due to isolated patches of not-fully ionized hydrogen produce the so-called Ly\(\alpha\) forest.

Recent improvements in observational facilities are boosting our ability to probe the tail end of the reionization process using QSO spectra. In particular, better spectral resolution and sensitivity, coupled with the discovery of QSOs at higher and higher redshift,\textsuperscript{7} are unveiling features at the edge of the reionization period. Barnett et al. (2017) presented a high-resolution spectrum of the QSO ULAS J1120+0641 at \(z_{\text{QSO}} = 7.084\) obtained with a 30 hr integration using the X-shooter instrument mounted on the Very Large Telescope. The only detected transmission blueward of the rest-frame Ly\(\alpha\) wavelength (with signal-to-noise ratio larger than 5) comes

\footnotesize{\textsuperscript{7}There are now more than 150 detected QSOs at \(z > 6.0\) according to Bosman (2019) and the current record holder is the QSO at \(z = 7.54\) discovered by Ballados et al. (2018).}
from seven narrow spikes in the Lyα forest. These span the redshift range $5.858 < z < 6.122$ and are located at the low-redshift edge of a long Gunn–Peterson trough (of size $\sim 240 \, h^{-1} \, \text{Mpc}$) extending all the way to the QSO proximity zone at $z_{\text{pq}} = 7.04$.

In a first attempt to extract information from similar observations, Gallarani et al. (2006, 2008) used a semianalytical model to relate the transmission windows and absorption gaps to the IGM ionized fraction, suggesting that transmission region are spatially associated with galaxies. If this result is confirmed, the radiation field should be enhanced in the regions where spikes are produced. We anticipate here, however, that such a geometrical argument does not directly entail a boosted transmission around galaxies because the sources of ionizing photons reside in overdense regions, where recombination is more efficient at suppressing the transmitted flux.

More recently, Chardin et al. (2018) combined the spectrum of ULAS J1120+0641 with a set of hydrodynamical simulations postprocessed with a radiative transfer code, showing that the number of spikes can, in principle, constrain the timing of reionization. Their work, however, relies on simulations of modest box sizes that neglect any coupling between gas dynamics and radiation, an effect that may be important for a detailed treatment of Lyα forest features.

In order to fully exploit available and forthcoming observations, in this paper we complement their analysis using three radiation-hydrodynamical simulations from the Cosmic Reionization On Computers (CROC) suite. We investigate the physical conditions and processes producing transmission spikes during the EoR and present the first detailed theoretical prediction for the correlation between spikes and neighboring galaxies. The text is organized as follows. After describing the simulation suite employed in this work, we present our main findings in Section 3, where we compare the simulations with a set of recent observations. Finally, we provide a summary and a final discussion of the implications of our results in Section 4.

2. Methods

Numerical studies of the EoR face formidable difficulties. The necessity to include a proper treatment of radiation transport (RT) increases the dimensionality of the problem and forces a compromise between simulation volume and resolution, but even after the ideal balance has been found, the comparison of simulated and real data is not straightforward because the simulated physical quantities are not directly observable. Hence, simulations have to be postprocessed in order to provide a meaningful comparison with observations. In this section, we describe the simulation suite employed in this work, as well as the production and subsequent characterization of synthetic absorption spectra.

2.1. CROC Simulations

In this work, we use three simulations from the CROC suite of Gnedin (2014). CROC is a set of radiation-hydrodynamical simulations performed using the Adaptive Refinement Tree (ART) code (Kravtsov et al. 1997). In particular, we employ simulations from the Caiman series, which include numerical convergence corrections (see Gnedin 2016) and a better time sampling of the tail end of reionization ($5 \lesssim z \lesssim 6.5$). We refer the interested reader to the CROC design paper (Gnedin 2014) and summarize in the following only the main features of the simulation set.

1. The heating and cooling rates of hydrogen and helium are computed self-consistently during the simulation run, without assuming photoionization or collisional equilibrium. In particular, the approximation of Gnedin & Hollon (2012) is used to determine the metallicity dependence of the heating and cooling functions as a function of the local radiation field.

2. Molecular hydrogen production and destruction is implemented using a fitting function calibrated against high-resolution self-consistent simulations (Gnedin & Draine 2014).

3. Star formation is included through an empirical subgrid model, following a linear Kennicut–Schmidt relation (Schmidt 1959; Kennicutt 1998) and assuming a typical star formation timescale of $\tau_{\text{SF}} = 1.5 \, \text{Gyr}$. The feedback from stars is implemented following the standard “delayed cooling” model (Stinson et al. 2006), while the ionizing radiation produced during their life is computed combining a Kroupa (2001) initial mass function with the Starburst99 model (Leitherer et al. 1999) for the spectral shape.

4. The CROC simulations are run employing the Optically Thin Variable Eddington Tensor (OTVET) approximation (Gnedin & Abel 2001, updated to suppress numerical diffusion) for the RT of UV (including ionizing) photons emitted by stars. Other sources of ionizing radiation (such as quasars, bremsstrahlung, and helium recombination radiation) are included as a background, either because they are weakly clustered or too rare (at the redshift covered by the simulations) to significantly influence a randomly selected region of the universe. The distortion of the background spectrum associated with the opacity of the optically thick Lyman-limit systems is computed employing the fit by Songaila & Cowie (2010).

5. The free parameters in the physical models have been calibrated against the observed galaxy UV luminosity function in the redshift range $5 \lesssim z \lesssim 10$ (that constrains the two parameters involved in star formation) and against the Gunn–Peterson optical depth of the Lyα forest in the spectrum of high-$z$ quasars (that constrains the ionizing photons escape fraction at the simulation resolution).

6. Density fluctuations on scales larger than the box size are treated using the “DC mode” formalism presented in Gnedin et al. (2011).

In this work, we employ three numerical simulations with box size $L_{\text{box}} = 80 \, h^{-1} \, \text{Mpc}$ (comoving). The computational box is initially discretized in a Cartesian grid containing $2048^3$ elements, and is subsequently adaptively refined to maintain spatial resolution approximately constant at 100 pc in physical units. At redshift $z = 6$, the corresponding Nyquist frequency $k_{\text{Ny}}$ (in velocity units) is $\approx 0.27 \, \text{s}^{-1}$ ($\pi/k_{\text{Ny}} \approx 12 \, \text{km} \, \text{s}^{-1}$), which is sufficient to model the IGM features in which we are interested. The simulations assume a present-day Hubble constant $H_0 = 68.14 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$, while the density parameters for baryonic matter, total matter, and cosmological constant are $\Omega_m = 0.3036$, $\Omega_b = 0.0479$, and $\Omega_\Lambda = 0.6964$. 
respectively (other cosmological parameters and simulation details are given in Gnedin 2014).

For the current analysis, we employ 13 redshift bins spanning the redshift range $5.0 \lesssim z \lesssim 6.5$. In each of these bins, simulation data were remapped onto a uniform, 1024$^3$ (cell size $= \pi/k_{ng}$) grid for the purpose of making synthetic absorption spectra. The three simulations we analyze follow the evolution of the same initial conditions but differ in their values of other cosmological parameters and simulation details.

Specifically, we impose $\delta_{\text{box}}$ to be $-1$, 0, and $+1$ times the theoretically expected rms density fluctuation in a cubic region of $80 h^{-1}$ Mpc on a side (i.e., $\delta_{\text{box}} = -\sigma_{\text{box}}$, 0, and $+\sigma_{\text{box}}$ respectively). This allows us to explore the effect of large-scale density fluctuations in a fully consistent way.

A visual impression of the simulations employed is given in Figure 1, where we show three different quantities across a slice through one simulation box: the baryonic overdensity $\delta_b \equiv \rho_b/\bar{\rho}_b - 1$ (with $\rho_b$ being the baryonic density field and $\bar{\rho}_b$ its value averaged over the entire simulation), the H I fraction $x_{\text{HI}}$, and the gas temperature $T$.

2.2. Synthetic Spectra

We produce synthetic H I Ly$\alpha$ absorption spectra by extracting the distribution of neutral hydrogen, gas temperature, and velocity along random lines of sight of length $L_s = 80 h^{-1}$ Mpc within the simulation box. The resolution elements of such spectra are $1 \text{km s}^{-1}$, to ensure that all simulated features are heavily oversampled and, hence, fully captured. Optical depths are transformed into (normalized) transmitted fluxes $f$ using a Voigt profile and including peculiar velocities and Doppler and thermal broadening. An example of a synthetic spectrum at redshift $z = 5.6$ is shown in Figure 2 (second row from the top). The baryonic overdensity surrounding the selected line of sight (dashed line) is reported in the top panel in a slice of size $24 h^{-1}$ Mpc $\times 1.8 h^{-1}$ Mpc of the simulation box. The remaining panels show the physical properties of the gas along the line of sight, namely (from middle to bottom) gas overdensity, H I fraction, and gas temperature.

At redshift $z \lesssim 4.5$, a substantial fraction of the IGM is transparent to Ly$\alpha$ photons, giving origin to the Ly$\alpha$ forest. At earlier times, however, most of the IGM is opaque to these photons, producing complete absorption everywhere but in a handful of highly ionized regions. The resulting spectrum shows only a few transmission spikes (see Figure 2). Quantitatively, we identify the latter as local maxima in the transmitted flux. The spike height ($h_s$) is defined as the maximum (normalized) transmitted flux, while the width ($w_s$) corresponds to the simply connected set of pixels that have $f \geq \alpha h_s$, where $\alpha$ is an adjustable parameter set to $\alpha = 0.5$ (Gnedin et al. 2017).

We associate each transmission spike to the physical properties of the cospatial IGM. In order to link a single value of each gas property to a spike, we employ as a representative value the average over the spike width weighted by the transmitted flux in each pixel, i.e.,

$$\bar{\delta} \equiv \frac{\int f(\lambda) q(\lambda) d\lambda}{\int f(\lambda) d\lambda},$$

where $q$ is a generic gas property and $\lambda$ is the wavelength. More specifically, in this work we associate to each spike the gas temperature $T$, baryonic overdensity $\delta_b$, and neutral fraction $x_{\text{HI}}$.

Similarly, we identify underdense regions along the line of sight. Starting from local minima in the overdensity ($\delta_{\text{min}}$), we define the underdense region as the (simply connected) region having $\delta < \beta\delta_{\text{min}}$, where $\beta$ is again an adjustable parameter. This provides us with two separate sets of line-of-sight portions that need to be matched. We do so by associating a spike to an underdense region whenever the pixel with the maximum transmitted flux is within the underdense region.
The second (third) panel from the top of Figure 2 show the results of these procedures. The shaded boxes highlight the spikes (underdense regions) identified in the spectrum. Each box spans vertically the region $[1, \alpha]$ ($[1, \beta]$) times the maximum (minimum) value and covers horizontally the full spike (underdense region) width. It appears clear that many underdense regions are not associated with any spike, while every spike is associated with an underdense region. This qualitative result (based on a single sight line) is investigated in more depth in the next section.

3. Results

In this section, we present the main findings of this work. We start by a careful physical characterization of the IGM underlying high-$z$ transmission spikes. Then, we move to study the average flux around bright sources, providing the first detailed predictions of such observable.

3.1. Spike Number Evolution

We start our analysis of the simulated spectra by characterizing the occurrence of transmission spikes as a function of redshift. They are identified as described in Section 2.2 and the resulting number $N_{\text{spikes}}$ per unit redshift per comoving megaparsec is shown in Figure 3 (solid lines) as a function of redshift. This number increases approximately linearly with redshift, driven by the decreasing average neutral fraction. We will investigate this correlation in more detail in Section 3.3.

Instrumental and observational effects, like finite spectral resolution and noise, reduce the measured value of $N_{\text{spikes}}(z)$. As an example, we simulate observations with the same resolution and noise, reduce the measured value of the former to $11\%$ of our synthetic spectra. This is shown in Figure 3 (dashed lines). Including these effects reduces the detected number of spikes by approximately an order of magnitude across all redshifts investigated. In the next section, we will use these “ULAS-like” spectra to determine in a quantitative way whether our simulations are in agreement with spikes observed in the spectrum of ULAS J1120+0641.

3.2. Comparison with ULAS J1120+0641

Before delving into the properties of high-$z$ transmission spikes, we assess whether our simulations are able to produce synthetic spectra similar to the spectrum of ULAS J1120+0641. To do so, we produce spectra of length $320 \ h^{-1} \ Mpc$ (comoving), approximately corresponding to the distance between the quasar proximity zone and the onset of the Ly$\alpha$ region. We do so by dividing the full $320 \ h^{-1} \ Mpc$ path length into eight segments (the number of simulation snapshots we have in this redshift interval) and concatenating randomly selected $40 \ h^{-1} \ Mpc$ long segments from consecutive simulation outputs at the appropriate redshift, assuming the quasar is located at $z_{\text{QSO}} = 7.084$. We thus account for light-cone effects, albeit only in a piece-wise constant approximation. We repeat this procedure 1000 times for each simulation box, for a total of 3000 spectra. In the redshift covered by the spectrum $(5.86 < z < 7.04)$, only few local features (transmission spikes) are present, and, therefore, the piece-wise constant approximation is not expected to introduce any artifact.

We show in Figure 4 one of these postprocessed synthetic spectra that is compatible with ULAS J1120+0641, together with the flux level $5\sigma$ above the noise. The latter is used as a threshold to identify spikes (i.e., spikes are required to have signal-to-noise ratio larger than 5). We consider our synthetic spectra compatible with that of ULAS J1120+0641 if they present $N_{\text{spikes}} = 7 \pm 3$ spikes in the first (i.e., lowest-redshift) $80 \ h^{-1} \ Mpc$ and $N_{\text{spikes}} \leq 1$ in the remaining $240 \ h^{-1} \ Mpc$. The uncertainties on these numbers are obtained as the Poisson error (rounded to the nearest integer) of the observed number of spikes in ULAS J1120+0641 for each of the aforementioned redshift ranges. We find that $\approx 15\%$ of our synthetic spectra are compatible with the one observed by Barnett et al. (2017). This number, however, varies significantly between simulations using a different DC mode, ranging from approximately 11% for the box with a positive DC mode, to 17% for the one with negative DC mode.

3.3. Transmission Spikes and IGM Properties

This paper is mainly devoted to the theoretical investigation of the connection between transmission region in the high-$z$ Ly$\alpha$ forest and the associated IGM. Therefore, we do not include noise and instrumental effects in the following analysis in order to provide a general view on this topic, which can be adjusted to both current and future observations. Additionally, the latter are still too sparse to allow a proper comparison. Nevertheless, we discuss in the Appendix how the results presented in this section are modified by the inclusion of observational effects.

In order to quantitatively study the physical connection between transmission spikes and the high-$z$ IGM, we start by investigating the 1D distribution of gas properties associated with transmission spikes in our synthetic spectra. This is shown in Figure 5 for three different redshift bins equally spaced in the range $z = 5.86, 6.16, 6.46$.
range $5.0 \leq z < 6.5$. All the spikes originate from regions that are underdense (i.e., $1 + \delta_b < 1$) at all the times investigated and become progressively emptier at higher redshifts. However, a low-density IGM is not sufficient to produce a spike, because the gas needs to be also highly ionized. The typical neutral fraction associated with spikes decreases from $x_{HI} \approx 3 \times 10^{-5}$ at $6.0 < z < 6.5$ to $x_{HI} \approx 1 \times 10^{-5}$ at $5.0 < z < 5.5$. At the redshift investigated, such combination of IGM properties is unlikely, explaining the rarity of the observed spikes. In particular, the production of these features requires underdense regions that experience a larger-than-average radiation field. To substantiate this statement, we compute for each spike the ratio between the associated neutral fraction ($x_{HI}^{\text{sim}}$) and the value expected from ionization equilibrium with the average background radiation at the density of the spike. The latter amount to $x_{HI}^{\text{eq.}} = x_{HI}(1 + \delta_b)$, where $x_{HI}$ is the average neutral fraction in the simulation box. The distribution of such ratios is shown in Figure 6 as a single histogram, because we checked that the three redshift bins give consistent results. The vast majority (more than 98%) of the spikes are produced in regions significantly more ionized than expected from ionization equilibrium with the average radiation field, i.e., $x_{HI}^{\text{sim}} / x_{HI}^{\text{eq.}} < 1$. We interpret this as a consequence of spikes...
production loci being preferentially close to a bright source of radiation. This is consistent with the conclusions from Gallerani et al. (2008), based on a semianalytical model, but we will investigate this interpretation in more details in Section 3.5.

Further information about the IGM and the sources of ionizing photons can be extracted from the shape of the spikes (Garaldi et al. 2019). Gnedin et al. (2017) showed that the CROC simulations produce a realistic distribution of spike heights and widths, with the exception of the $w < 800 \text{ km s}^{-1}$ tail of the spike width distribution. We are, therefore, confident that the simulations employed in this work are suitable for our purposes. In Figure 7 we plot the joint distributions of spike widths and heights with gas overdensity, temperature, and H I fraction. The distributions show that the variables investigated are mostly uncorrelated, therefore making it hard to translate the shape of a spike into direct information about the underlying gas. Nevertheless, the very observation of a spike at a given redshift can be used constrain the underlying IGM physical properties. While these can cover a broad range at $5.0 < z < 5.5$, at even earlier epochs our results indicate that only a relatively narrow range of IGM overdensities underlies the observed spikes.

The spike height appears to be broadly correlated with the gas density, especially at lower redshifts, and to the underlying neutral fraction. In particular, $h_s$ is more sensitive to $x_{\text{HI}}$ in the redshift range $5.5 < z < 6.0$; in this interval, the contours are more tilted with respect to the axes, while in the other two redshift ranges, the contours are almost parallel to either the horizontal (at $6.0 < z < 6.5$) or the vertical (at $5.0 < z < 5.5$) axis. The distribution of IGM temperatures is largely insensitive to spike width and height, as well as redshift.

### 3.4. Spike Width and Underdense Regions

Figure 7 shows that at any given redshift, the width of the spikes is not correlated with any of the IGM properties investigated. This naturally raises the question of what physical property sets these widths. We have shown that spikes are produced exclusively in underdense, overionized regions of the IGM. Hence, it appears likely that their width is simply determined by the extent of such regions along the line of sight. In order to investigate this hypothesis, we identify underdense regions along the line of sight and match them with spikes as described in Section 2.2. We associate an underdensity width ($w_{ur}$) to each spike and investigate the connection between $w_s$ and $w_{ur}$. The joint distribution of these two quantities is shown in Figure 8 for spikes that have an associated underdense region, color-coded with respect to the number density of spikes in any given pixel. For ease of comparison, we express both $w_s$ and $w_{ur}$ in $\text{km s}^{-1}$. We measure the former from spectra
and we converted their physical length in velocity space. In the case of underdense regions, we have investigated the effect of this parameter on the discussed correlation and show results for the value of $\beta$ that produce an approximate 1:1 relation in the lowest-redshift bin, i.e., $\beta = 0.8$. We stress, however, that the physical interpretation of this result does not change, because varying this parameter has the only effect of changing $w_{ur}$ at fixed $w_{s}$.

For completeness, we report in the bottom panel the distribution of $w_{ur}$ associated with spikes in the three redshift bins. Its peak shifts toward larger value with increasing redshift, a trend that can be understood as a consequence of the larger underdensities required to produce transmission spikes at earlier times (see Figure 5).

3.5. Spike–Galaxy Correlation

We have shown in Section 3.3 that transmission spikes are produced in underdense regions that are overionized. We interpreted this combination of physical properties as an indication that the production loci of these features are close to a UV source while still being underdense relative to the background. Similar conclusions have been obtained recently by Kakiichi et al. (2018, K18 hereafter), who investigated observationally the concurrence of transmission spikes and neighboring galaxies by means of a spectroscopic campaign. Their results, although limited to a single line of sight, additionally show that the relation between spikes and nearby galaxies can be used as a tool to understand the sources of reionization and their properties. It is likely, as well as desirable, that this kind of study will soon involve a much larger number of lines of sight to high-$z$ quasars. Therefore, it is of key importance to theoretically investigate this link in order to guide future observational efforts, as well as to provide theoretical predictions that can be contrasted with forthcoming data. In the following, we employ the CROC simulations to this end.

We provide a first visual impression of the concurrence of galaxies and transmission spikes in Figure 9, where we show a synthetic Ly$\alpha$ spectrum extracted from our simulations and the galaxies within 4 Mpc from the line of sight. The galaxy color coding reflects their UV magnitude (computed from simulated stellar particles using the Flexible Stellar Population Synthesis library Conroy et al. 2009; Conroy & Gunn 2010), and only objects brighter than the K18 detection threshold are shown. This figure is, effectively, a synthetic version of Figure 7 in K18. Also depicted in the figure is the IGM baryonic overdensity along the line of sight, clearly indicating that all the observed spikes coincide with large underdense regions. Note how the presence of a nearby bright galaxy is not a sufficient condition for the production of a spike. Similarly, many underdense regions are not associated with a transmission spike.

In order to quantify the concurrence of transmission spikes and nearby galaxies, K18 performed a correlation analysis, evaluating the average flux transmitted as a function of distance.

Figure 8. Joint distribution of spike widths ($w_{s}$) and underdense region lengths along the line of sight ($w_{ur}$) in simulated spectra within three different redshift bins (top three panels). The color coding indicates the density of points in each pixel. Red lines show the best-fit linear relation between the two quantities. The bottom panel depicts the 1D normalized distribution of the sizes of underdense region in the same redshift bins.
from nearby (spectroscopically detected) galaxies:

$$\langle f(r) \rangle = \frac{1}{N_{\text{pair}(r)}} \sum_{i \in \text{pair}(r)} \hat{f}_i.$$  \hspace{1cm} (3)

Here, $N_{\text{pair}(r)}$ is the total number of pairs in the ensemble of spectral pixel—galaxy pairs at a given distance $r$, i.e.,

$$\text{pair}(r) \equiv \{ p \in \mathcal{P}_{\text{LOS}}, g \in \mathcal{G} \} \text{ dist}(p, g) = r,$$  \hspace{1cm} (4)

where $\mathcal{P}_{\text{LOS}}$ is the ensemble of pixels in all available spectra (only one in the case of K18), $\mathcal{G}$ is the collection of all spectroscopically detected galaxies, and dist$(p, g)$ denotes the physical distance between a pixel $p \in \mathcal{P}_{\text{LOS}}$ and a galaxy $g \in \mathcal{G}$.

We have computed the same quantity from our numerical simulations and compared it with the data and the linear-theory predictions of K18. There are, however, a few improvements with respect to their analysis: (i) we have repeated the procedure at seven different redshifts ($z = 5.0, 5.2, 5.4, 5.6, 5.9, 6.1,$ and $6.4$), (ii) we have averaged the results over $10^3$ random spectra to build each curve, and (iii) we have included all galaxies resolved in our simulations, which corresponds to a much lower UV detection threshold—$M_{UV}^{\text{min}} \sim -16$. For completeness, we report here that the results presented below are not affected by the latter choice, as increasing the minimum magnitude up to $\sim -21.2$ does not change the results, although it decreases their statistics. We stress, however, that $M_{UV}^{\text{min}}$ is just a postprocessing parameter and is not linked in any way to the minimum magnitude of galaxies producing ionizing photons in the simulation.

In the left panel of Figure 10 we show $\langle f(r) \rangle$ extracted from our simulations. Similar to the prediction from K18, there is a $1\text{pMpc}$ region where $\langle f(r) \rangle$ increases steadily until it saturates at exactly the average flux within the entire box, losing any information about the pixel—galaxy distance. The...
drop in the average flux transmitted at $r \lesssim 1$ pMpc is consistent with being due to the effect of the overdensities that host the galaxies (see, e.g., Appendix B in K18). We have checked that this is indeed the case in our simulation by studying the average neutral hydrogen fraction as a function of $r$. This quantity shows a radial dependence complementary to the one of $\langle f(r) \rangle$, i.e., a sharp increase in the inner $\sim 1$ pMpc followed by a leveling off at its average value in the box. Interestingly, the physical size of the region with suppressed flux (defined as the distance where $f(r)$ is 95% of the mean value) remains approximately constant with redshift (in the relatively narrow range investigated). Unlike K18, however, we do not find any region at intermediate distances where $f(r)$ is enhanced with respect to the box mean. We will comment again on this discrepancy at the end of this paragraph, after presenting other relevant observations. In addition, we note here that computing $\langle f(r) \rangle$ separately for galaxies of different stellar mass does not change the result. This indicates that, in our simulation, the enhanced flux coming (on average) from larger galaxies is balanced by the larger density such galaxies reside in (on average, due to the tight stellar mass—halo mass relation, e.g., Moster et al. 2010).

The shape of the curves $\langle f(r) \rangle$ is consistent across all the simulations, i.e., for different values of $\delta_{\text{box}}$. However, their vertical amplitude is strongly dependent on the latter. In particular, the simulation with a large-scale positive (negative) density fluctuation $\delta_{\text{box}} = + \sigma_{\text{box}} (- \sigma_{\text{box}})$ shows a strongly suppressed (enhanced) average flux with respect to the box with $\delta_{\text{box}} = 0$. The $\langle f(r) \rangle$ in the latter is very close to the mean of the three simulations. This effect can be seen in the top right panel of Figure 10, which we shall discuss below. Before moving on, however, we note here that the main cause of the differences between simulations with various DC modes is in the timing of structure formation and evolution. The offset between the different curves is, in large part, due to one of them ($\delta_{\text{box}} > 0$) evolving faster than the average universe and the other ($\delta_{\text{box}} < 0$) evolving slower than average. We have checked that this is indeed the case by comparing $\langle f(r) \rangle$ in different runs at the same value of $\sigma_{\text{box}}$ (i.e., the expansion factor of the box, which is different than the universal scale factor as a consequence of the different mean density), and indeed the three boxes produce almost indistinguishable results.

In Figure 10 we compare our results to recent high-redshift observations by K18 (top panel) and Meyer et al. (2019, bottom panel). The measurements in these two studies are presented in different forms and cover slightly different epochs, and we therefore discuss them separately. To compare with the data of K18, we select from each simulation the snapshot at the epoch that is the closest to the average redshift of the galaxies detected by K18, i.e., $z \approx 5.8$. In the right panel of the figure, we show the results from each simulation box separately, in order to display the effect of large-scale density fluctuations. The agreement of the simulated $\langle f(r) \rangle$ with the observations depends strongly on the DC mode of the box. The best overall match is provided by $\delta_{\text{box}} = 0$, in particular when taking into account the fact that the increased observed $\langle f(r) \rangle$ at $3 \lesssim r/\text{pMpc} \lesssim 5$ is an artifact due to the resampling of a single, prominent transmission spike (K18). It should be noted, however, that the data of K18 are based on a single line of sight and, therefore, can potentially be biased. Hence, additional observations are necessary to allow a statistically significant statements about the ability of current reionization models to reproduce the observed average flux. Our results also show, however, that this task is hindered by the fact that a proper comparison between synthetic and real data requires knowledge of the large-scale overdensity of the observed field, which is mostly unavailable, especially at the high redshifts were Ly$\alpha$ transmission spikes are detected. Therefore, to effectively compare theoretical predictions to observed spectra, the asymptotic value of $\langle f(r) \rangle$ needs to be factored out. This is done e.g., in Meyer et al. (2019), where the normalized transmission is employed in place of $\langle f(r) \rangle$. We note here that the dependence on the DC mode can be used as a powerful tool to determine the large-scale density field from a relatively narrow patch of the sky. In fact, by comparing the flux at $r \gtrsim 2.5$ pMpc with its simulated value as a function of box overdensity, we can infer the latter from observations. In particular, applying this reasoning to the observations by K18 suggests that the observed field is close to the average density of the universe on large scales.

Meyer et al. (2019) performed a study similar to K18, but using CIV absorbers as tracers of low-mass galaxies and focusing on a slightly lower-redshift interval centered around $z \approx 5.4$. We show their results (obtained combining 25 different quasar lines of sight) in the bottom panel of Figure 10. Different from K18, they normalize the transmitted flux to the average along each line of sight. This procedure effectively erases differences due to large-scale fluctuations. In fact, when we use our simulations to produce a synthetic version of such quantity, we obtain consistent results across the different simulation boxes. In particular, the outcome of our simulations is in good agreement with the observational data. Simulations, however, appear to slightly underpredict the observed values at intermediate ($2 \lesssim r/\text{pMpc} \lesssim 6$) distances. At first, this result may appear at odds with the larger-than-average ionization field experienced by the gas producing the transmission spikes (Figure 6). However, the boosted ionization field around sources is balanced by an increased neutral fraction. Therefore, the presence of (a region of) boosted transmission entirely depends on the balance between these two opposing effects, which may vary with the distance from the central galaxy.

Finally, we shall discuss here the discrepancy between our results and the linear-theory-based predictions of K18, which show a proximity zone of enhanced $\langle f(r) \rangle$ at $r \gtrsim 2$ $h^{-1}$ Mpc before approaching its asymptotic ($r \to \infty$) value at larger radii. As detailed above, this may hint to a different balance between recombinations and photoionization at intermediate radii. In the model of K18, this boost is a consequence of the galaxy placed at $r = 0$ that enhances the surrounding flux thanks to its radiation field. However, in deriving such a prediction, K18 assumed that the galaxy located at $r = 0$ is the only detected galaxy (and hence the brightest) in the range of $r$ explored. We can test whether this assumption holds in our simulations. The faintest spectroscopically detected galaxy in K18 has $M_{\text{UV}}^{\text{th}} = -21.1$. In our simulation, galaxies above this mass threshold have, on average, $N_{\text{neigh}} = 5$ galaxies with $M_{\text{UV}} \geq M_{\text{thr}}^{\text{th}}$ within 6 pMpc. Therefore, the region $r < 6$ pMpc is affected by the proximity effect of multiple galaxies, enhancing the $\langle f(r) \rangle$ with respect to the prediction obtained assuming a single bright source. The number $N_{\text{neigh}}$ rapidly increases with decreasing $M_{\text{thr}}^{\text{th}}$, rising to $N_{\text{neigh}} = 21.4$ for $M_{\text{thr}}^{\text{th}} = 10^{9} h^{-1} M_{\odot}$. Supporting this view is also the fact that
the asymptotic value of $\langle f(r) \rangle$ predicted by the model of K18 is much lower than what is obtained in our simulations, possibly indicating that this value is due to an overlap of the contribution of multiple sources.

There are, however, two possible alternative explanations. First, if our reionization model overestimates the galactic escape fraction, then the proximity region of a galaxy extends to larger radii than in reality and, at the same time, the $\langle f(r) \rangle$ profile becomes flatter at much shorter distances (see the bottom left panel of Figure 8 in K18). This could indeed be the case, because the UV escape fraction at the simulation resolution ($\epsilon_{\text{UV}}$) is a free parameter that has been calibrated to reproduce the optical depth distribution observed in the Ly$\alpha$ forest. The value employed, $\epsilon_{\text{UV}} = 0.15$, is in good agreement with numerical simulations of escape of ionizing radiation from molecular clouds (Howard et al. 2017, 2018) at similar spatial scales (100 pc), but there exist no observational constraints on this value at present.

The second possibility is that our simulations underestimate the minimum magnitude of galaxies contributing to the ionizing photon budget. In this case, unresolved galaxies produce enough ionizing photons to increase the asymptotic $\langle f(r) \rangle$ and suppress the proximity effect of the bright $r = 0$ galaxy (see the bottom right panel of Figure 8 in K18).

It is likely that a combination of all these effects is driving the difference observed in $\langle f(r) \rangle$. Unfortunately, it is not easy to disentangle individual contributions in numerical simulations. In particular, it is highly nontrivial to predict the ionizing state of the IGM with different contributions from smaller galaxies. Similarly, changing the escape fraction will change the IGM ionization state and history in a way that is unpredictable at the level of detail required for the present study. Therefore, if the inclusion of nearby bright sources in the theoretical prediction of K18 or future observations will not solve the discrepancy, a large suite of numerical simulations with different values of relevant parameters is needed to assess which is the main reason for the difference between the analytical prediction of K18 and our results. At the moment this appears—unfortunately—prohibitively expensive.

### 4. Summary and Conclusions

In this work, we have highlighted the information that can be extracted from the Ly$\alpha$ transmission regions (“spikes”) embedded in the extended Gunn–Peterson troughs observed in high-redshift quasar spectra. We have used state-of-the-art CROC radiation-hydrodynamic simulations to produce synthetic spectra including the effects of gas peculiar velocity, thermal and Doppler broadening, finite spectral resolution, and instrumental noise. We employed a simple algorithm to identify spikes in the normalized transmitted flux and underdense regions along the line of sight (Figure 2). Our results are based on three runs sharing the same initial conditions but different values of the density fluctuation on the scale of the box (the DC mode). Our results can be summarized as follows.

1. The number of transmission spikes as a function of redshift evolves in the simulations from 0 at $z \gtrsim 6$ to approximately 11 Mpc$^{-3}$ at $z = 5$. At the resolution and signal-to-noise ratio level of current observations of the distant quasar ULAS J1120+0641, this number is reduced by an order of magnitude at all redshifts. The number of spikes is larger (at all redshifts) in the run with a positive DC mode, while it is smaller when the DC mode is negative. This is a consequence of the faster (slower) evolution of the former (latter) with respect to a box having exactly the mean density of the universe.

2. Synthetic versions of a ULAS J1120+0641-like spectrum obtained by concatenating simulation outputs at different redshifts show good agreement with the data in approximately 15% of the cases.

3. Spikes are found to originate exclusively from underdense, overly ionized regions that become less underdense and more ionized with time. The inclusion of observational effects removes the high-density tail of the IGM underdensity distribution as the associated spikes have, on the average, a lower transmitted flux.

4. The spike width does not correlate with the IGM density at any given redshift. The spike height is negatively correlated with the gas (over)density, but with a large scatter and in a redshift-dependent fashion. The height also correlates with the gas ionized fraction, more so at $z \approx 5.75$. Including observational effects removes the correlation between spike height and gas density, as a consequence of a preferential suppression of high-density, small-height spikes. The correlation with the gas ionized fraction is mostly unchanged (Figure 12).

5. Transmission spikes at redshift $5.0 \leq z \leq 6.5$ are produced in regions that are more ionized than expected from ionization equilibrium with the average UV background flux. Spikes are, therefore, formed in underdense regions of the IGM that are sufficiently close to a source of ionizing radiation.

6. The spike width is determined by the extent of the associated underdense region along the line of sight. The link between width and extent changes with time (for a fixed width parameter $\beta$).

7. The average transmission in synthetic absorption spectra as a function of distance from a bright galaxy shows a proximity zone of suppressed flux with an approximately constant radius in physical coordinates. The zone is associated with the overdense galaxy region. The average transmission at any given redshift depends on the value of the DC mode. Recent observations of the transmission as a function of source distance by Kakiichi et al. (2018) and Meyer et al. (2019) show some mild indications of the existence of a region with enhanced flux at intermediate radii, which is missing in our simulations. We identified two possible shortcomings of our simulations: the overprediction of the escape fraction or the underestimation of the minimum magnitude of galaxies that significantly contribute to the reionization of the universe.

Our analysis shows that high-redshift transmission spikes are a promising tool to push investigations of the IGM well into the tail end of cosmic reionization. We predict such spikes to be ubiquitous in quasar sight lines and estimate that only approximately 10% of all predicted spikes are detectable even with the best data currently available. Once more of these spectral features are discovered, the way forward will ideally follow two parallel paths. On the one hand, the larger statistics will promote spikes to a powerful and complementary probe of reionization models in an era where the average Ly$\alpha$ optical depth is already very large (Chardin et al. 2018; Kakiichi et al. 2018). On the other hand, detailed simulations that will prove able to predict the distribution and shape of the observed spikes...
should be used to draw a physical connection between such spectral features and the ionization and thermal state of the early IGM.

We thank Koki Kakiichi for useful discussions on transmission spikes. We also thank the anonymous referee for their comments. Support for this work was provided by NASA to P.M. through a contract to the WFIRST-EXPO Science Investigation Team (No. 15-WFIRST15-0004), administered by the GSFC. This research has been partially carried out within the SFB 956 “The Conditions and Impact of Star Formation,” subproject C4, funded by the Deutsche Forschungsgemeinschaft (DFG). E.G. thanks the University of California, Santa Cruz for their warm hospitality, during which this work was conceived and initiated. Fermilab is operated by Fermi Research Alliance, LLC, under contract No. DE-AC02-07CH11359 with the United States Department of Energy. This work was partly supported by a NASA ATP grant No. NNX17AK65G, and used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under contract No. DE-AC02-06CH11357. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research is also part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards Nos. OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. We made extensive use of the NASA Astrophysics Data System and arXiv.org preprint server, and are thankful to the community developing and maintaining the software packages Matplotlib (Hunter 2007), NumPy (Walt et al. 2011), SciPy (Jones et al. 2001).

Appendix

Simulating ULAS J1120+0641

Many of the results discussed in this work are instrument-agnostic, based on synthetic spectra that have a very high spectral resolution and no noise. This has allowed us to produce an unbiased view of the physical mechanisms at the origin of the Lyα transmission spikes and, at the same time, to produce products that can easily be translated in predictions for any given combination of spectral resolution and noise level. Here, we assess the impact of realistic observational effects using the observations of ULAS J1120+0641 (Barnett et al. 2017) as a template (see Section 3.1). In particular, we present here the “observer version” of Figures 5 and 7.

Figure 11 shows the properties of the intergalactic gas responsible for the transmission spikes. A comparison with its idealized counterpart (Figure 5) reveals that observational effects preferentially erase spikes originating from the less-underdense regions. This is a consequence of finite spectral resolution smearing out narrow spikes and reducing their height. The latter is smaller in less-underdense regions (see Figure 7), which are therefore preferentially suppressed once a noise threshold is imposed.

Figure 12 is the “observer version” of Figure 7 (notice that the axes have different ranges, for the sake of visual clarity). The inclusion of realistic observational conditions has a two-fold impact. The first is the shrinking of the allowed properties for the gas producing detectable spikes. The second is due to finite spectral resolution broadening the spikes and stretching the range of spike widths by more than a factor of 5. These effects also induce a cutoff in the minimum $w_s$ at approximately 200 $\text{km s}^{-1}$, which is larger than the spectral resolution ($\Delta v = 10 \text{ km s}^{-1}$). Resolution suppresses the flux of narrow spikes as they are smeared over a larger spectral range, so that only broad spikes survive the noise threshold selection. An exception to this condition is represented by spikes that are very narrow but also very high, so that even after being smeared out they are still detectable above the noise. Such spikes are rare and therefore are not contained in the 95% contour shown in the figure. We note here that observational effects do not change significantly the results of Figure 6, although they reduce the overall number of detectable spikes as already discussed.
References
Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, Natur, 553, 473
Barnett, R., Warren, S. J., Becker, G. D., et al. 2017, A&A, 601, A16
Bosman, S. 2019, All z > 5.7 Quasars Known at the Moment, http://www.
sarahbosman.co.uk/list_of_all_quasars.htm
Chardin, J., Haeblen, M. G., Bosman, S. E. L., & Puchwein, E. 2018, MNRAS, 473, 765
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018, ApJ, 864, 142
Eide, M. B., Graziani, L., Ciardi, B., et al. 2018, MNRAS, 476, 1174
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117
Gallerani, S., Choudhury, T. R., & Ferrara, A. 2006, MNRAS, 370, 1401
Gallerani, S., Ferrara, A., Fan, X., & Choudhury, T. R. 2008, MNRAS, 386, 359
Garaldi, E., Compostella, M., & Porciani, C. 2019, MNRAS, 483, 5301
Gnedin, N. Y. 2000, ApJ, 542, 535
Gnedin, N. Y. 2014, ApJ, 793, 29
Gnedin, N. Y. 2016, ApJ, 821, 50
Gnedin, N. Y., & Abel, T. 2001, NewA, 6, 437
Gnedin, N. Y., Becker, G. D., & Fan, X. 2017, ApJ, 841, 26

Figure 11. Same as Figure 5, but for ULAS J1120+0641-like spectral resolution and noise.

Figure 12. Same as Figure 7, but for ULAS J1120+0641-like spectral resolution and noise.
Gnedin, N. Y., & Draine, B. T. 2014, ApJ, 795, 37
Gnedin, N. Y., & Hollon, N. 2012, ApJS, 202, 13
Gnedin, N. Y., Kravtsov, A. V., & Rudd, D. H. 2011, ApJS, 194, 46
Haardt, F., & Madau, P. 1996, ApJ, 461, 20
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Howard, C., Pudritz, R., & Klessen, R. 2017, ApJ, 834, 40
Howard, C. S., Pudritz, R. E., Harris, W. E., & Klessen, R. S. 2018, MNRAS, 475, 3121
Hunter, J. D. 2007, CSE, 9, 90
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/
Kakiichi, K., Ellis, R. S., Laporte, N., et al. 2018, MNRAS, 479, 43
Kennicutt, R. C., Jr. 1998, ApJ, 498, 541
Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, ApJS, 111, 73
Kroupa, P. 2001, MNRAS, 322, 231
Kulkarni, G., Worseck, G., & Hennawi, J. F. 2018, arXiv:1807.09774
Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3
Madau, P., & Haardt, F. 2015, ApJL, 813, L8
Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648
Marchi, F., Pentericci, L., Guaita, L., et al. 2017, A&A, 601, A73
Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, ApJ, 856, 2
McGreer, I. D., Mesinger, A., & Fan, X. 2011, MNRAS, 415, 3237
Meyer, R. A., Bosman, S. E. I., Kakiichi, K., & Ellis, R. S. 2019, MNRAS, 483, 19
Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903
Ota, K., Iye, M., Kashikawa, N., et al. 2010, ApJ, 722, 803
Pen, U.-L. 1997, ApJL, 490, L127
Pentericci, L., Fontana, A., Vanzella, E., et al. 2011, ApJ, 743, 132
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv:1807.06209
Schmidt, M. 1959, ApJ, 129, 243
Schroeder, J., Mesinger, A., & Haiman, Z. 2013, MNRAS, 428, 3058
Sirko, E. 2005, ApJ, 634, 728
Songaila, A., & Cowie, L. L. 2010, ApJ, 721, 1448
Stinson, G., Seth, A., Katz, N., et al. 2006, MNRAS, 373, 1074
Vanzella, E., Nonino, M., Cupani, G., et al. 2018, MNRAS, 476, L15
Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22