New methods for identifying Lyman continuum leakers and reionization-epoch analogues

Harley Katz,[1]⋆ Dominika Ďurovčíková [ID],[2] Taysun Kimm [ID],[3] Joki Rosdahl [ID],[4] Jeremy Blaizot,[4] Martin G. Haehnelt,[5] Julien Devriendt,[1] Adrianne Slyz,[1] Richard Ellis[6] and Nicolas Laporte [ID][5,7]

[1]*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[2]*New College, University of Oxford, Holywell Street, Oxford OX1 3BN, UK*
[3]*Department of Astronomy, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea*
[4]*Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France*
[5]*Kavli Institute for Cosmology and Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK*
[6]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*
[7]*Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK*

ABSTRACT

Identifying low-redshift galaxies that emit Lyman continuum radiation (LyC leakers) is one of the primary, indirect methods of studying galaxy formation in the epoch of reionization. However, not only has it proved challenging to identify such systems, it also remains uncertain whether the low-redshift LyC leakers are truly 'analogues' of the sources that reionized the Universe. Here, we use high-resolution cosmological radiation hydrodynamics simulations to examine whether simulated galaxies in the epoch of reionization share similar emission line properties to observed LyC leakers at $z \sim 3$ and $z \sim 0$. We find that the simulated galaxies with high LyC escape fractions ($f_{\rm esc}$) often exhibit high O32 and populate the same regions of the R23–O32 plane as $z \sim 3$ LyC leakers. However, we show that viewing angle, metallicity, and ionization parameter can all impact where a galaxy resides on the O32–$f_{\rm esc}$ plane. Based on emission line diagnostics and how they correlate with $f_{\rm esc}$, lower metallicity LyC leakers at $z \sim 3$ appear to be good analogues of reionization-era galaxies. In contrast, local [S II]-deficient galaxies do not overlap with the simulated high-redshift LyC leakers on the S II Baldwin–Phillips–Terlevich (BPT) diagram; however, this diagnostic may still be useful for identifying leakers. We use our simulated galaxies to develop multiple new diagnostics to identify LyC leakers using infrared and nebular emission lines. We show that our model using only [C II]$_{158\,\mu{\rm m}}$ and [O III]$_{88\,\mu{\rm m}}$ can identify potential leakers from non-leakers from the local Dwarf Galaxy Survey. Finally, we apply this diagnostic to known high-redshift galaxies and find that MACS 1149_JD1 at $z = 9.1$ is the most likely galaxy to be actively contributing to the reionization of the Universe.

Key words: HII regions – galaxies: high-redshift – galaxies: ISM – galaxies: star formation – dark ages, reionization, first stars.

1 INTRODUCTION

While it is well established that the Universe transitioned from a nearly neutral to an ionized state during its early evolution, much remains uncertain about this process of reionization. Reionization likely began at $z \gtrsim 30$ with the formation of the first Population III stars (Wise 2019). However, the redshift at which reionization ended is currently debated, with estimates ranging from $z \sim 6$ (Fan, Carilli & Keating 2006) to $z \sim 5.3$ (Kulkarni et al. 2019). In addition to the history of reionization, the sources that provided the photons that ionized the intergalactic medium (IGM) remain unknown. Numerous possibilities exist including dwarf galaxies and minihaloes (Couchman & Rees 1986; Bouwens et al. 2015), massive galaxies (Naidu et al. 2020), active galactic nuclei (AGN; Haiman & Loeb 1998; Madau & Haardt 2015), accretion shocks (Dopita et al. 2011), globular clusters (Ricotti 2002; Katz & Ricotti 2013, 2014), stellar mass black holes (Madau et al. 2004; Ricotti & Ostriker 2004;

Mirabel et al. 2011), and dark matter annihilation and decay (Mapelli, Ferrara & Pierpaoli 2006).

Ideally, one could directly observe these objects during the epoch of reionization and simply measure the amount of Lyman continuum (LyC) radiation that escapes the source and ionizes the IGM. However, in practice, this is incredibly difficult, if not impossible, because of the Gunn–Peterson effect (Gunn & Peterson 1965), which causes the absorption of nearly all photons blueward of Lyα due to a neutral IGM. Hence our understanding of the escape of ionizing photons from their source populations into the IGM during the epoch of reionization comes from predominantly two methods: (1) using numerical radiation hydrodynamics simulations to simulate the process from first principles (Wise et al. 2014; Kimm et al. 2017; Trebitsch et al. 2017; Rosdahl et al. 2018), and (2) using observations to study low-redshift 'analogues' that are either leaking LyC photons or inferred to be leaking LyC radiation that may be representative of the high-redshift source population (e.g. Borthakur et al. 2014; Siana et al. 2015; Vanzella et al. 2015; Leitherer et al. 2016; Shapley et al. 2016; Marchi et al. 2017; Matthee et al. 2017; Naidu et al. 2018).

⋆ E-mail: harley.katz@physics.ox.ac.uk

Despite substantial differences in modelling approaches, certain trends have emerged from high-resolution numerical simulations regarding how LyC radiation escapes galaxies. The general consensus is that this process is 'feedback regulated', indicating that strong supernova (SN), or perhaps AGN feedback, is required to clear channels from the centres of galaxies through which LyC radiation can escape (Kimm et al. 2017; Trebitsch et al. 2017, 2020). Galaxies are expected to evolve through 'bursty' phases whereby the galaxy is most often opaque to LyC radiation and this only changes after a substantial star formation episode. Simulations also tend to show that observing the escaping radiation is viewing angle dependent (Cen & Kimm 2015). Furthermore, galaxies of different masses do not leak LyC radiation equally and simulations tend to agree that low-mass galaxies are much more efficient in this regard (Wise et al. 2014; Paardekooper, Khochfar & Dalla Vecchia 2015; Xu et al. 2016; Kimm et al. 2017), although see e.g. Ma et al. (2020).

Observationally, identifying low-redshift systems with leaking LyC radiation has proved a difficult exercise, with most estimates of the LyC escape fraction (f_{esc}) being less than what is required to reionize the Universe (Siana et al. 2007, 2010; Bridge et al. 2010; Leitet et al. 2011, 2013; Leitherer et al. 2016; Rutkowski et al. 2016). Knowing where to look is one of the key challenges and there has been considerable success at $z \gtrsim 3$ (Steidel et al. 2018; Fletcher et al. 2019; Meštrić et al. 2020; Nakajima et al. 2020). Low-redshift studies often focus on low-metallicity star-forming galaxies that have similar properties to higher redshift Lyα emitters (e.g. Dressler et al. 2015; Trainor et al. 2015; Nakajima et al. 2016). Many of the leakers that have been observed can be characterized by certain emission line properties such as high line ratio of O32 ($\log_{10}$([O III] 4960, 5007 Å/[O II] 3727, 3729 Å); de Barros et al. 2016; Faisst 2016; Izotov et al. 2018b; Nakajima et al. 2020), a small separation between the red and blue Lyα peaks (Verhamme et al. 2017; Izotov et al. 2018b), or deficits in S II emission (Wang et al. 2019). Many of these trends can be understood using idealized models for H II regions (Guseva et al. 2004; Jaskot & Oey 2013; Nakajima & Ouchi 2014) and by studying the differences between density-bounded and ionization-bounded H II regions. Nevertheless, it still remains to be determined whether the systems with observed LyC leakage at $z \lesssim 6$ are actually representative of the high-redshift galaxies that may be reionizing the Universe.

In this work, we compare the emission line properties of low-redshift LyC leakers with those of simulated galaxies that are forming deep into the epoch of reionization to determine whether low-redshift 'analogues' are representative of the high-redshift galaxy population. Furthermore, we use the emission line properties of our simulated galaxies to develop new methods for identifying LyC leaker 'analogues' at low redshift. Throughout this work, we define LyC leakers from non-leakers as galaxies that have an $f_{esc} \geq 10$ per cent. This separation is motivated by the fact that in order to complete reionization by $z \sim 6$, one needs to adopt an escape fraction for galaxies in the range of ~ 7–20 per cent (Robertson et al. 2015; Rosdahl et al. 2018).

This paper is organized as follows. In Section 2, we introduce the ASPEN simulations and how we model the LyC escape fraction and emission line luminosities. In Sections 3 and 4, we compare our simulated systems with observations and derive new diagnostics for identifying LyC leakers from observations. We test these diagnostics against observations at $z = 0$ and $z > 6$ to identify potential local LyC leakers and galaxies that played an active role in reionizing the Universe. Finally, in Sections 5 and 6, we present our discussion and conclusions.

2 NUMERICAL METHODS

In this section, we describe our cosmological radiation hydrodynamics simulations and how we measure the escape fractions and emission lines from each of the simulated galaxies.

2.1 Cosmological radiation hydrodynamics simulations

For this work, we employ the state-of-the-art ASPEN suite of simulations (Katz et al. 2019a,b), a set of cosmological radiation hydrodynamics simulations that zoom-in on the formation of a massive galaxy in the epoch of reionization. Complete details of these simulations are presented in Katz et al. (2019b) and here, we briefly summarize the main properties of the simulations.

Our simulation zooms in on the formation of a Lyman-break galaxy that has a dark matter halo mass of $10^{11.8}$ $M_\odot$ by $z = 6$. Initial conditions were generated for both dark matter and gas with MUSIC (Hahn & Abel 2011) at $z = 150$ using a convex-hull shape for the high-resolution region. We employ a Planck Collaboration XIII (2016) cosmology with the following cosmological parameters: $h = 0.6731$, $\Omega_m = 0.315$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.685$, $\sigma_8 = 0.829$, and $n_s = 0.9655$. The effective resolution of our simulation is 4096^3, corresponding to a dark matter particle mass of 4×10^4 $M_\odot$ h^{-1}. We use the Amiga Halo Finder (AHF; Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009) to extract haloes from our simulation and we define the virial radius to be that which contains a mean density of dark matter, gas, and stars equal to $\Delta \rho_{crit}$, where ρ_{crit} is the critical density of the Universe and Δ is the overdensity that would allow for spherical collapse in an expanding cosmological background at a given redshift. We consider only galaxies deep into the epoch of reionization where $\Delta \sim 200$. Furthermore, we only study haloes that are completely uncontaminated by low-resolution dark matter particles and that are sampled by a minimum of 300 dark matter particles. This corresponds to a minimum dark matter halo mass in our simulation of 1.2×10^7 $M_\odot$ h^{-1}, which is well below the atom cooling threshold mass. By $z = 10$, there are more than 1000 uncontaminated haloes of at least this mass. We assume an initial gas composition in the simulation of 76 per cent H and 24 per cent He by mass.

Gravity, hydrodynamics, radiative transfer, and non-equilibrium chemistry are modelled using the version of RAMSES (Teyssier 2002; Rosdahl et al. 2013; Rosdahl & Teyssier 2015) presented in Katz et al. (2017) and Kimm et al. (2017). For the hydrodynamics, we use the MUSCL-Hancock scheme with a Harten–Lax–van Leer contact (HLLC) Riemann solver (Toro, Spruce & Speares 1994) and a MinMod slope limiter (Roe 1986). The relation between gas pressure and internal energy is closed by assuming the gas is ideal and monatomic with $\gamma = 5/3$. Radiative transfer is solved with a first-order moment method [using the M1 closure (Levermore 1984) and a GLF intercell flux function (Toro 2009)]. We parametrize the radiation field with eight frequency bins from infrared (IR) to He II ionizing that are coupled to the gas thermochemistry via photoionization, photoheating, and radiation pressure [both direct ultraviolet (UV) and multiscattered IR]. Non-equilibrium chemistry is solved for seven species: H I, He I, H II, He II, He III, H_2, and e^-. To make the radiative transfer computational tractable, we use the reduced speed of light approximation and set the speed of light in the simulation to be $0.01c$. Furthermore, we subcycle the radiative transfer calculation up to 500 times per hydrodynamic time step (Rosdahl et al. 2013; Commerçon, Debout & Teyssier 2014). We model the cooling of the gas in a non-equilibrium manner for the species listed above taking into account collisional ionizations, recombinations, collisional excitation, bremsstrahlung, Compton

cooling and heating, and dielectronic recombination (see appendix E of Rosdahl et al. 2013). H_2 cooling is included using the rates from Hollenbach & McKee (1979) as are equilibrium rates for metal line cooling (Rosen & Bregman 1995).

Star formation is modelled in the simulation based on a thermoturbulent recipe (Kimm et al. 2017; Rosdahl et al. 2018) and a Schmidt law (Schmidt 1959). The stellar mass in our simulations is set to a minimum of 1000 $M_\odot$. Star particles inject radiation into their host cell based on their mass, age, and metallicity, according to a BPASS v2.0 spectral energy distribution (SED; Eldridge, Izzard & Tout 2008; Stanway, Eldridge & Becker 2016) assuming an initial mass function (IMF) with a maximum stellar mass of 300 $M_\odot$ and a slope of -1.30 between 0.1 and 0.5 $M_\odot$ and -2.35 between 0.5 and 300 $M_\odot$. For the first 50 Myr during the lifetime of each star particle, they can undergo SN that is randomly drawn from a realistic delay-time distribution (Kimm et al. 2015). SNe are modelled using the mechanical scheme of Kimm et al. (2015) and the equivalent of 10^{51} erg is injected back into the gas for each SN. In the case where the H II region around a star particle is unresolved, a fixed value of 5×10^5 $M_\odot$ km s^{-1} of momentum per SN is injected (Geen et al. 2015; Kimm et al. 2017). For each SN, 20 per cent of the mass of the star is recycled back into the gas assuming it is metal enriched to a value of 0.075 (Kroupa 2001). We employ a $4\times$ boost to the number of SN, consistent with Rosdahl et al. (2018) to match the high-redshift extrapolation of the stellar mass–halo mass relation. This also produces a reasonable reionization history and UV luminosity function (Rosdahl et al. 2018).

Because we use an adaptive mesh refinement (AMR) technique for our simulation, we use a quasi-Lagrangian technique to refine cells. When a cell contains either eight dark matter particles or the gas mass of the cell is $> 8 \frac{\Omega_b}{\Omega_{DM}} m_{DM}$, where m_{DM} is the dark matter particle mass, the cell is refined. Similarly, a cell is also refined if the cell width is more than 1/4 of the local Jeans length. We allow the simulation to refine to maintain a constant physical resolution of 13.6 pc.

For full details of the simulation, refer to Katz et al. (2019a,b).

2.2 Nebular and IR emission line luminosities

In order to compare with observations, we need to measure different emission line luminosities for each galaxy in our simulation. We do this in post-processing by employing the photoionization code CLOUDY (Ferland et al. 2017). For each cell in the simulation, we know the temperature, density, metallicity, and local radiation field (strength and shape). Ideally, we would run a CLOUDY model for each cell; however, this is currently computationally unfeasible due to the total number of cells in the simulation and the number of snapshots that we wish to post-process. We estimate that this would require at least a few million CPU hours. For this reason, we employ a random forest (Ho 1995; Breiman 2001; Geurts, Ernst & Wehenkel 2006), which is an ensemble machine learning technique to quickly estimate the various emission line luminosities that we are interested in. The random forest was trained on 850 000 cells that were extracted from the $z = 10$ snapshot of the simulation, and in Katz et al. (2019b), we demonstrated that by using our trained model, we can estimate the total luminosities of galaxies in our simulation to better than 5–10 per cent accuracy for most luminous lines[1] at a fraction of the computational cost of running CLOUDY. Full details of the CLOUDY

models and their accuracies are presented in Katz et al. (2019b). For this work, we will only study the intrinsic line luminosities for each galaxy (not including dust attenuation), measured by summing the luminosity of each gas cell within the virial radius of each halo. Furthermore, we only consider IR and nebular emission lines and will not study resonant lines (such as Lyα), although they are indeed very interesting in the context of estimating the escape fraction (Verhamme et al. 2015, 2017; Dijkstra, Gronke & Venkatesan 2016; Kimm et al. 2019).

2.3 Measuring the escape fraction

To measure the escape fraction (f_{esc}) of galaxies in our simulations we post-process the simulations with a ray-tracing method. Methods such as these have been shown to be equally good as directly measuring the ionizing flux emanating from the galaxies (Trebitsch et al. 2017). For each galaxy in the simulation, we extract a sphere around the centre of each halo with radius equal to the virial radius of the system, and we project the neutral hydrogen density in each cell on to a cube (with a side length equal to the diameter of the halo) that has a cell width equal to the maximum physical resolution of the simulation in that snapshot ($\sim$13.6 pc). For each star particle in the halo, we then ray trace (Buchner & Bauer 2017) through the cube in 500 random directions from the location of each star particle, keeping the ionization state of the gas fixed, to measure the optical depth to neutral hydrogen and thus the escape fraction. We only consider neutral hydrogen here and we assume a mean photon energy of 21 eV, which is the mean energy of a photon across the age- and metallicity-dependent SEDs that we employ, up to a metallicity of 0.5 $Z_\odot$. We then take a luminosity-weighted average of each star particle to measure the total f_{esc} of each galaxy. Filling the grid with dust, molecular hydrogen, and helium and ray tracing at multiple photon energies does not substantially change our measured f_{esc} and our method is significantly computationally cheaper (see fig. 5 of Kimm et al. 2019). We have also checked that 500 rays are sufficient by measuring the differences between using between 100 and 100 000 rays.
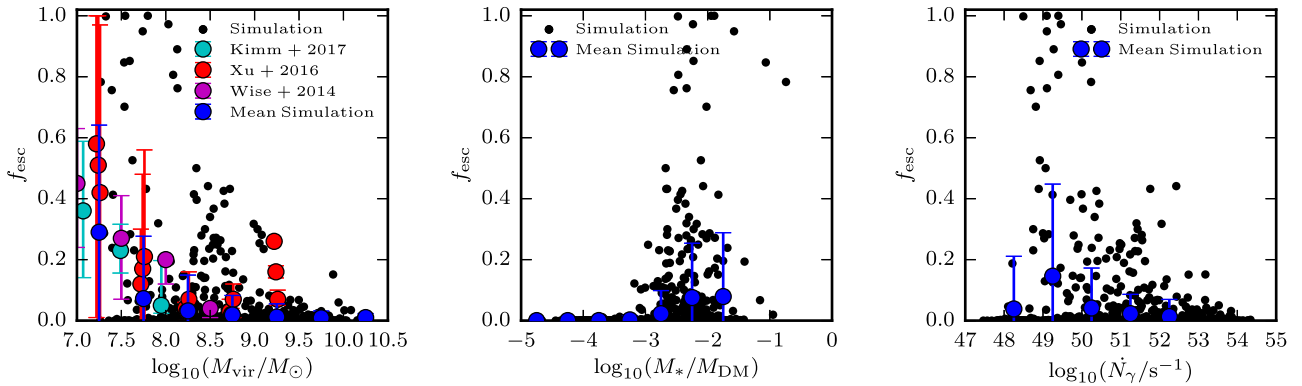
We measure f_{esc} for all galaxies in our simulation that host at least 20 star particles (i.e. that have a minimum stellar mass of 2×10^4 $M_\odot$). Furthermore, in this work, we will only consider the global f_{esc} (i.e. that averaged over all directions) so that it can be compared to the intrinsic line luminosities. It should be noted that in observations, f_{esc} can change quite substantially depending on viewing angle (Cen & Kimm 2015). However, our method is more useful for determining how the physics of leaking radiation is related to emission line luminosities, without the complications of inhomogeneous dust and neutral hydrogen distributions. We note that if the distribution of dust is significantly different from that of the neutral hydrogen in real galaxies, observations will be impacted in a different way from the simulations where the dust is defined to be in the metal enriched, mostly neutral gas.

3 RESULTS

In this section, we discuss how the results from our simulations compare to others in the literature and then use our simulations to deduce whether observed galaxies that have either high value of O32 or [S II] deficit are analogues of reionization-era galaxies.

3.1 Global properties

In this work, we aim to elucidate whether the properties of low-redshift LyC leakers are consistent with those during the epoch of

---

[1]This accuracy is already better than the systematic uncertainty of changing the geometry of the calculation (e.g. switching from a slab to a sphere that can change the results on the order of a few per cent).

Figure 1. Left: virial mass of the galaxy versus LyC escape fraction at $12 \geq z \geq 9.2$. There are nine simulation snapshots stacked in each of these plots. Individual galaxies in our simulation are shown as black points, while blue points with error bars represent mean values of f_{esc} within bins of 0.5 dex in virial mass. For comparison we show the results from simulations of Kimm et al. (2017) in cyan, Xu et al. (2016) in red, and Wise et al. (2014) in purple. Despite differences in star formation routines, feedback schemes, and spatial and mass resolution, good agreement is seen among the different simulations. Centre: ratio of stellar mass to halo mass of the galaxy versus LyC escape fraction for individual galaxies in the simulation. Right: instantaneous, intrinsic ionizing luminosity versus LyC escape fraction for individual galaxies in the simulation.

reionization and derive diagnostics for identifying the LyC leakers at low redshift. Using the ASPEN simulations, we have analysed the properties of all galaxies at nine different redshifts, deep into the epoch of reionization, in the range $12 \geq z \geq 9.2$. In Fig. 1, we plot f_{esc} versus the virial mass of the halo for all galaxies in our simulation at all of the sampled redshifts. Consistent with previous works (Wise et al. 2014; Xu et al. 2016; Kimm et al. 2017), we identify a strong trend of f_{esc} with mass such that the lower mass systems exhibit higher average escape fractions compared to higher mass objects. The individual black points represent individual galaxies in our simulation, while the blue points represent the mean values in bins of virial mass with a width of 0.5 dex. Despite differences in feedback models, resolution, and redshifts analysed, our escape fractions are consistent with the trend in other work in the literature (Wise et al. 2014; Xu et al. 2016; Kimm et al. 2017).

Simulations now tend to agree that the escape of LyC photons is feedback regulated and that a strong burst of star formation combined with effective SN feedback is required for such photons to escape (e.g. Kimm & Cen 2014; Kimm et al. 2017; Trebitsch et al. 2017; Rosdahl et al. 2018; Ma et al. 2020). One may expect that the galaxies that have a high ratio of stellar mass to dark matter mass are the ones that likely have strong enough feedback to disrupt the gas in the centre of the halo and create channels for LyC photons to escape. In the centre panel of Fig. 1, we plot the ratio of M_*/M_{DM} versus f_{esc}. In general, only the haloes with higher M_*/M_{DM} exhibit significant LyC leakage; however, we note that this could be impacted by simulation resolution as the higher resolution simulations of Kimm et al. (2017) demonstrate that certain galaxies with low M_*/M_{DM} can exhibit significant LyC leakage. Lower spatial resolution can smooth the interstellar medium (ISM), which may inhibit LyC photons from escaping. Not all systems with high M_*/M_{DM} exhibit high f_{esc} as the timing of when the stars formed also matters. Since f_{esc} fluctuates, the time period to catch the phase of high f_{esc} after a star formation event can be as short as 1–10 Myr (see e.g. Kimm et al. 2017; Trebitsch et al. 2017; Yoo, Kimm & Rosdahl 2020) and thus not all systems with high M_*/M_{DM} can be expected to have high f_{esc} at the fixed redshifts that we analyse. This timing argument also applies to the lowest mass systems with low M_*/M_{DM}. These galaxies have the fewest number of star particles and thus the probability of catching them in an optically thin phase is low.
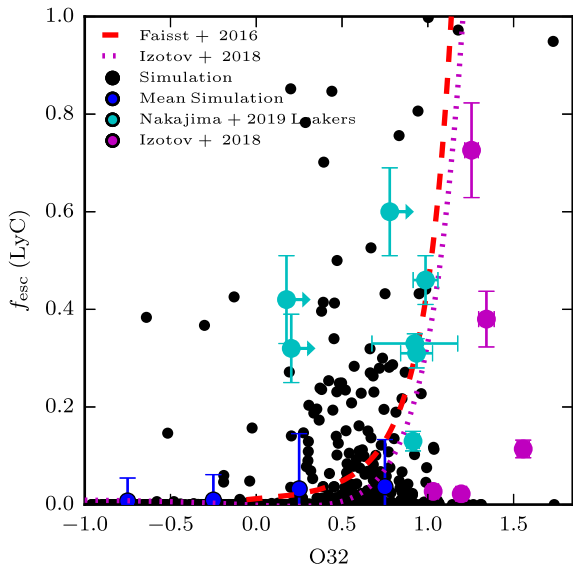
In the right-hand panel of Fig. 1, we compare the instantaneous production rate of ionizing photons ($\dot{N}_\gamma$) with f_{esc}. There is considerable scatter in this plot; however, high escape fractions are preferentially found in less luminous galaxies in LyC. This is reflective of the fact that the lower mass objects in our simulations tend to have higher f_{esc} as there is a strong trend between virial mass and f_{esc}. Certain highly luminous systems also have high f_{esc}, many of which represent lower mass galaxies that have undergone a recent burst of star formation.

The global properties of the galaxies in our simulation seem to agree well with previous works (Wise et al. 2014; Xu et al. 2016; Kimm et al. 2017), both in the context of their escape fraction properties and their star formation and emission line properties (Katz et al. 2019b). In subsequent sections, we relate the observational properties of our simulated galaxies with those of low-redshift LyC leakers, with an emphasis on nebular and IR emission lines.

3.2 O32

One method that has been used for identifying LyC leakers at low redshift is to focus on systems that have excess [O III] emission compared to [O II] emission (O32, $\log_{10}$([O III] 4960, 5007 Å/[O II] 3727, 3729 Å); e.g. Nakajima & Ouchi 2014). From a theoretical viewpoint, O32 correlates with ionization parameter, the ratio of ionizing photon density to hydrogen density (Kewley & Dopita 2002), although there is an additional dependency on metallicity. Galaxies with high-ionization parameters are more likely to be leaking ionizing photons, hence O32 is expected to be a useful diagnostic for identifying LyC leakers. Many observational works have confirmed that galaxies that exhibit leaking LyC radiation often have high O32 (Nakajima & Ouchi 2014; de Barros et al. 2016; Faisst 2016; Izotov et al. 2018a; Nakajima et al. 2020).

In Fig. 2, we plot O32 versus f_{esc} for our simulated galaxies compared with low-redshift observations. Consistent with the low-redshift observations, we find that the simulated galaxies that exhibit a significant fraction of LyC leakage also tend to be biased towards high O32. All of the observed galaxies that have high f_{esc} have O32 > 0. If we use O32 = 0 as our dividing line between high and low O32, we find that in our sample of simulated galaxies, 11 per cent of systems with high O32 have f_{esc} > 10 per cent. In contrast, only 1 per cent of the simulated galaxies with O32 < 0 have

Figure 2. O32 ($\log_{10}$([O III] 4960, 5007 Å/[O II] 3727, 3729 Å)) versus f_{esc} for all galaxies in our simulation (black points). Cyan and purple points represent observed low-redshift LyC leakers from Nakajima et al. (2020) and Izotov et al. (2018b), respectively. The blue dashed and purple dotted lines show the expected trends between O32 and f_{esc} from Faisst (2016) and Izotov et al. (2018b), respectively. In general, we find that systems that have a significant amount of LyC leakage also exhibit higher O32.

$f_{esc} > 10$ per cent. The trend lines from Faisst (2016) and Izotov et al. (2018b) represent reasonable approximations of our simulated galaxies although they tend to be slightly biased towards higher O32 for a given f_{esc}. However, it is unclear if these trend lines match the simulations at the highest values of O32 where the simulated galaxies are sparsely sampled. There is considerable scatter in this relation expected from both the observations and our simulations and it is also important to note that our work employs the angle averaged quantities for these values while observations can only probe a single line of sight.

Recently, Bassett et al. (2019) have called into question the utility of this diagnostic as not all observed galaxies with high O32 also have high f_{esc}. It is clear from Fig. 2 that this is indeed the case in our simulations. Nakajima et al. (2020) have also identified a sample of observed galaxies at $z \sim 3$ that have high O32 and low f_{esc} and attribute this to the impact of viewing angle. In contrast, Bassett et al. (2019) outline numerous idealized models of H II regions that show how the relation between O32 and f_{esc} responds to variations in metallicity and ionization parameter. The results from our simulations demonstrate that LyC leakers are preferentially biased towards having high O32 and below, we study both effects outlined in the literature.

3.2.1 Impact of viewing angle

Nakajima et al. (2020) recently studied a sample of Lyα emitters at $z \sim 3.1$ and found that although the galaxies leaking LyC radiation were biased towards higher O32, this was not a sufficient condition for the galaxy to be leaking radiation. Because there were no other identified spectral differences between the galaxies that were leakers versus non-leakers, they suggested that viewing angle was a possible explanation.

In the top panel of Fig. 3, we show a Mollweide projection of neutral hydrogen column density for a galaxy that is leaking LyC

photons ($f_{esc} = 0.15$) and that has a higher O32 compared to many other systems in our sample (O32 = 0.52). This halo has a dark matter mass of $10^{9.9}$ M$_\odot$ and a stellar mass of $10^{7.9}$ M$_\odot$. We define n_{clear} to be the fraction of viewing angles where an observer would measure a $f_{esc} > 10$ per cent and for this system, we find $n_{clear} = 38$ per cent. We would expect that most of the time, an observer would measure a high O32 but a low f_{esc}. This figure demonstrates that for a given galaxy, the measured LyC f_{esc} is strongly dependent on viewing angle (see also Cen & Kimm 2015). Hence, for observed galaxies, viewing angle can significantly impact the trend between O32 and f_{esc}, consistent with the arguments made in Nakajima et al. (2020).
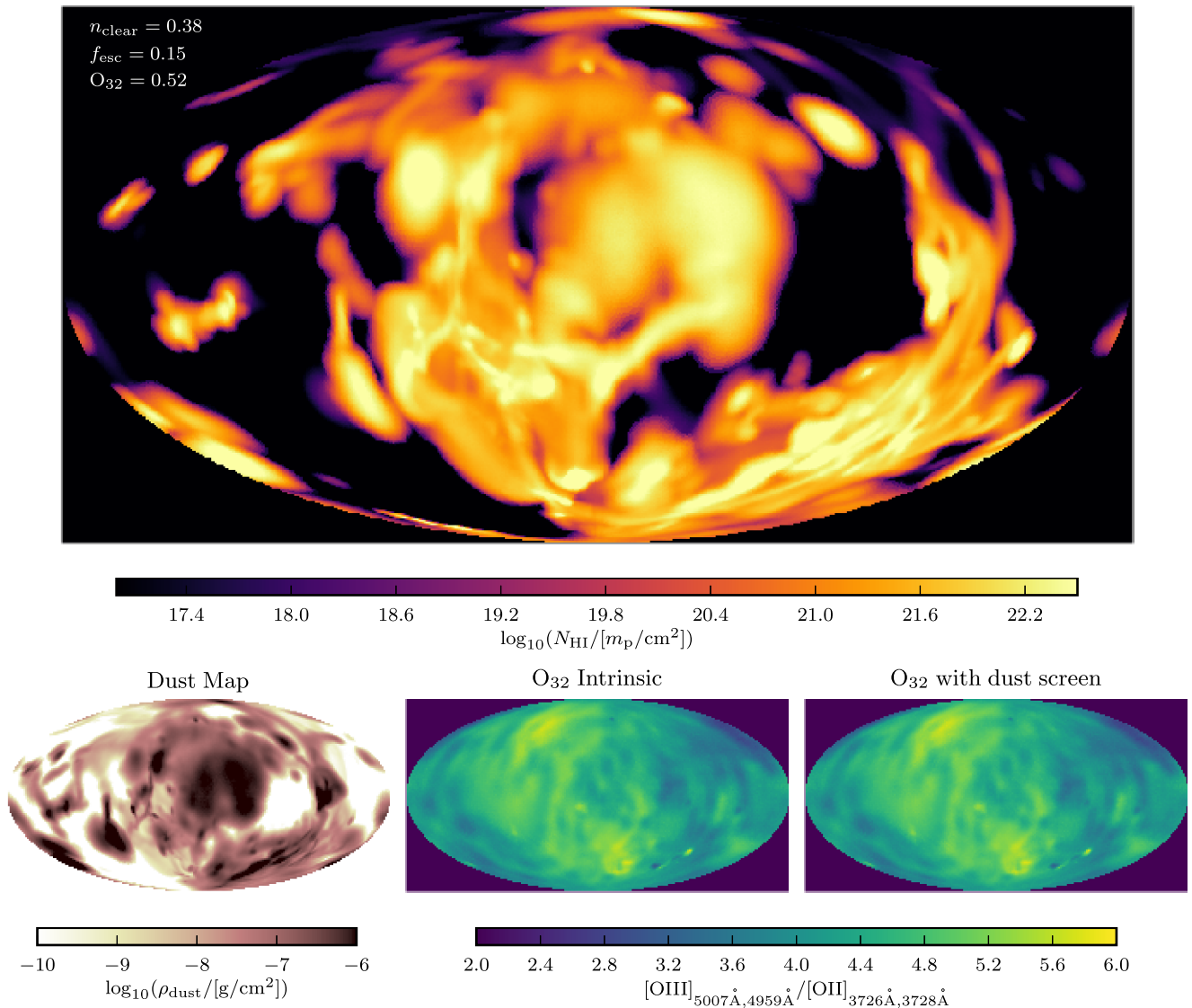
In addition to angular variations in f_{esc}, it is plausible that there are also angular variations in O32. According to Bassett et al. (2019), dust is not expected to significantly impact O32, with corrections expected to be <0.3 dex. In the bottom left-hand panel of Fig. 3, we show a Mollweide projection of the dust column density. The dust mass in each simulation cell is computed following the description in Katz et al. (2019a) where we apply a metallicity-dependent dust-to-metal ratio (Rémy-Ruyer et al. 2014) and assume that all dust has been destroyed in cells that have $T > 10^5$ K. The dust distribution is clearly very inhomogeneous and tends to follow the location of neutral hydrogen. In the bottom centre panel of Fig. 3, we show line-of-sight variations in O32 (in linear scale) that differ by a factor of 3. In the bottom right-hand panel, we show the dust attenuated O32 Mollweide projection (assuming that the inhomogeneous dust map acts as a dust screen for each line)[2] and we only find ~5 per cent differences accounting for dust. Most of the leakers in our simulation have low metallicity so it is unlikely that for this work, viewing angle will significantly impact O32. In contrast, much higher metallicity galaxies could see a larger effect as indicated in Bassett et al. (2019).

3.2.2 Impact of metallicity and ionization parameter

While viewing angle can certainly impact the observed relation between O32 and f_{esc}, from Fig. 2, it is clear that there is a population of high-redshift galaxies that exhibit high O32 but low f_{esc} that are not viewing angle dependent. This is because, for this figure, we show the angle-averaged f_{esc}, rather than that along a specific line of sight. This opens the question of what intrinsic properties cause galaxies to have high O32 but low f_{esc}?

In Fig. 4, we show examples of two galaxies, one with high O32 and high f_{esc} (top) and the other with high O32 but low f_{esc} (bottom). The properties of these two galaxies are listed in Table 1. We have selected these two massive systems in haloes with mass >10^9 M$_\odot$ so that they are sufficiently resolved by our simulation. For each galaxy, we show the mass-weighted temperature–density phase diagram, a map of gas column density, a map of stellar surface mass density, maps of [O III]$_{4960\,Å\,+\,5007\,Å}$ and [O II]$_{3727,\,3729\,Å}$ intensity, and a map of the O32 spatial distribution, from left to right. There are clear and obvious structural differences between the two systems. The system with high O32 and high f_{esc} has both an extended gas and stellar distribution and there is clear evidence that stellar feedback and other dynamical effects have disturbed the structure of the galaxy. This is evident in the phase diagram where significant amounts of ionized gas are present and likewise there is a considerable amount of gas within the virial radius that has a low density and a high temperature, a consequence of effective supernova feedback. In contrast, the system with high O32 and low f_{esc} is very compact, is

---

[2]Assuming that the dust distribution acts as a screen would represent the maximum impact of dust on O32.
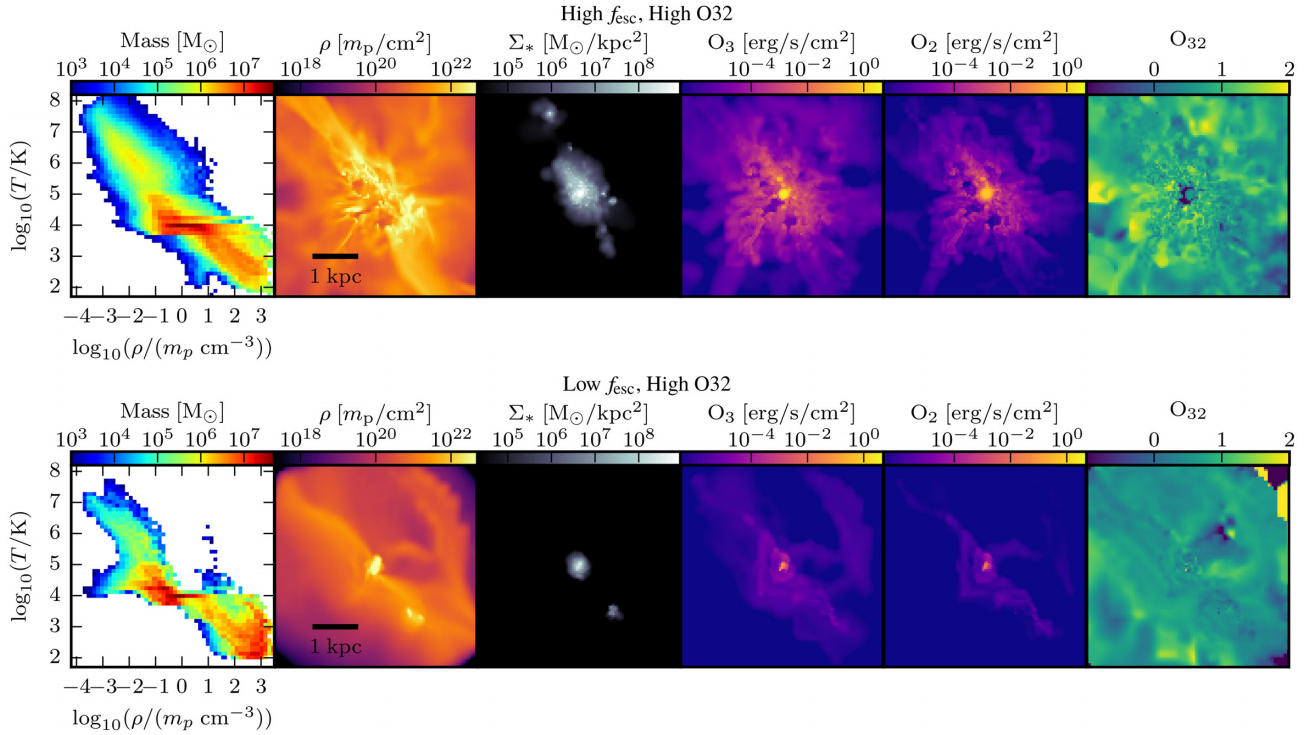
Figure 3. Top: Mollweide projection of neutral hydrogen column density as observed from the centre of a galaxy with high O32 and leaking LyC radiation. This galaxy has a dark matter mass of $10^{9.9}$ M$_\odot$ and a stellar mass of $10^{7.9}$ M$_\odot$. $n_{\rm clear}$ represents the fraction of sightlines that have an escape fraction of at least 10 per cent. Depending on the viewing angle, an observer can either view an escape fraction of 0 or up to 100 per cent. Bottom: Mollweide projections of the dust column density (left), intrinsic O32 (centre), and O32 accounting for dust (right). The dust tends to follow the locations of neutral hydrogen, while O32 variations are more sensitive to the strength of the radiation field and feedback along a specific line of sight. O32 is shown in linear scale on these images and different lines of sight show differences of a factor of 3 in O32. This galaxy is low metallicity and even though the dust distribution is highly inhomogeneous, the impact on O32 is only $\sim$5 per cent.

not sufficiently impacted by stellar feedback, as seen from the phase diagram, and has very little ionized gas. Much of this behaviour is driven by the fact that the first galaxy has a considerably higher stellar mass. Although usually discussed in the context of H II regions, from a galaxy perspective, the first system with high O32 and high $f_{\rm esc}$ can be thought of as a density-bounded galaxy (i.e. when there is not enough gas to absorb all of the ionizing radiation), while the second system with high O32 and low $f_{\rm esc}$ can be thought of as an ionization-bounded galaxy (i.e. when the source of radiation fills out its Stromgren sphere inside the cloud).

The system with low $f_{\rm esc}$ has about half the metallicity and an order of magnitude lower stellar mass compared to the galaxy with high $f_{\rm esc}$. From the photoionization models of Bassett et al. (2019, see the right-hand panel of their fig. 10), it is clear that

for fixed ionization parameter and $f_{\rm esc}$, low-metallicity systems have systematically higher O32, likely caused by low-metallicity SEDs being harder. This also causes our simulated systems to fall significantly higher on the R23–O32 (R23 = $\log_{10}(([\text{O III}]\ 4960,\ 5007\ \text{Å} + [\text{O II}]\ 3727,\ 3729\ \text{Å})/\text{H}\beta)$) diagram compared to Sloan Digital Sky Survey (SDSS) galaxies (see Fig. 5), yet still consistent with observed systems at $z \sim 3$ (Strom et al. 2017; Katz et al. 2019b; Nakajima et al. 2020). One can see from Fig. 5 that this plane is not particularly useful for differentiating between LyC leakers and non-leakers. Furthermore, in the models of Bassett et al. (2019), for fixed ionization parameter, the upturn in $f_{\rm esc}$ as a function of O32 occurs at systematically higher O32 as metallicity decreases. The galaxies in our simulation have varying metallicities and in our two example galaxies, the trends with $f_{\rm esc}$ follow exactly that predicted

Figure 4. From left to right: mass-weighted temperature–density phase-space diagram, gas column density map, stellar surface mass density map, O3 intensity map, O2 intensity map, and spatial distribution of O32. The top panel shows an example of a galaxy with high O32 and high f_{esc} (Halo 1), while the bottom panel shows and example of a galaxy with high O32 and low f_{esc} (Halo 2). The properties of Halo 1 and Halo 2 can be found in Table 1. Individual regions of a galaxy might have very high or very low O32 depending on local conditions and deviate strongly from the value averaged across the galaxy. Despite both having high O32, structural, metallicity, stellar population, and feedback differences between the two systems mean that they have very different LyC escape fractions.

Table 1. Properties of the haloes shown in Fig. 4.

Halo	f_{esc} (%)	M_{halo} (M$_\odot$)	M_* (M$_\odot$)	SFR (M$_\odot$ yr^{-1})	Metallicity (Z$_\odot$)	O32
Halo 1	15.1	7.66×10^9	7.64×10^7	0.24	0.030	0.52
Halo 2	0.0014	1.76×10^9	6.57×10^6	0.08	0.017	0.67

by Bassett et al. (2019). It is important to emphasize that ionization parameter (and also strength of stellar feedback) also matters in determining f_{esc}.
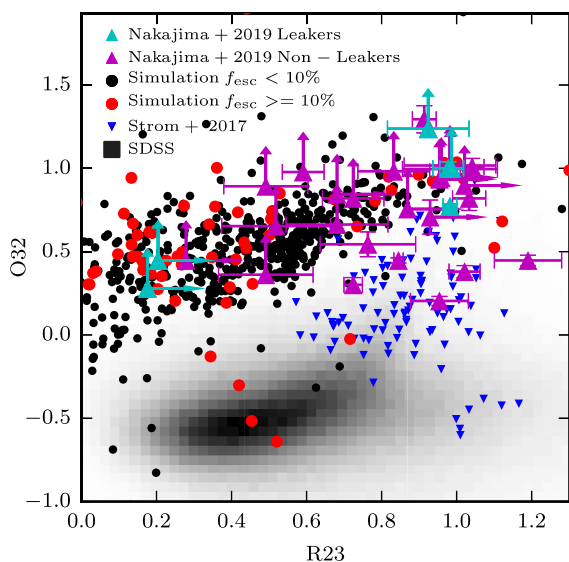
In the top panel of Fig. 6, we show O32 versus f_{esc} where we have colour coded the points by the metallicity of the galaxy. In the middle panel of Fig. 6, we plot O32 versus galaxy metallicity colour coded by ionizing luminosity. From the middle panel, we see that the systems at fixed O32 with higher metallicity also tend to have higher ionizing luminosities. This is a result of higher metallicity galaxies being biased towards having higher stellar masses. In the bottom panel of Fig. 6, we show O32 versus mean H I ionization rate inside the virial radius colour coded by f_{esc}. As expected, we find that the galaxies with higher O32 are biased towards higher ionization parameter. However, there is considerable scatter due to the degeneracy with metallicity and ionization parameter. While it is difficult to identify trends in O32–f_{esc} plane, there is a hint that the galaxies in our simulation with the highest O32 (e.g. O32 > 1) have lower metallicities and higher H I ionization rates.

In summary, the relationship between O32 and f_{esc} is complex and dictated by quantities such as metallicity, ionization parameter, morphology, and feedback. The majority of the galaxies in our simulations with high f_{esc} are biased towards having high O32,

consistent with observations (Izotov et al. 2018a; Nakajima et al. 2020). Quantitatively, the majority of our LyC leakers have O32 > 0 and all of the observed LyC leakers have O32 greater than this value. Using O32 = 0 as a dividing line, if we normalize to the number of galaxies in both populations only 1 per cent of the galaxy population with O32 < 0 have f_{esc} > 10 per cent. In contrast 11 per cent of galaxies with O32 > 0 have f_{esc} > 10 per cent. Hence, galaxies with higher O32 are more biased to be LyC leakers. However, consistent with Bassett et al. (2019), despite this bias, we do not find any other strong trends that may help identify more LyC leakers based on f_{esc} and O32 alone. Thus for isolated systems, high O32 may be a very common, but not sufficient condition for LyC leakage. It is important to note that Bassett et al. (2019) also identified two systems with high f_{esc} but low O32. These two systems show evidence of merging and the shocks from this merger may decrease O32.

3.3 S II deficit

Recently, observations from Wang et al. (2019) suggested that [S II] emission can be used to potentially identify LyC leakers at low redshift. They demonstrated that LyC leakers exhibited systematically lower [S II] emission on the [S II] Baldwin–Phillips–Terlevich (BPT)
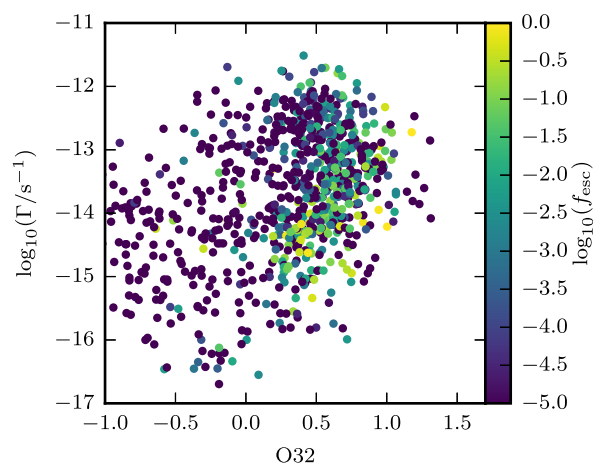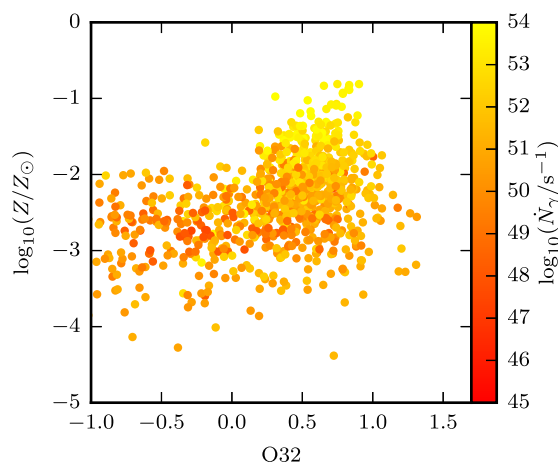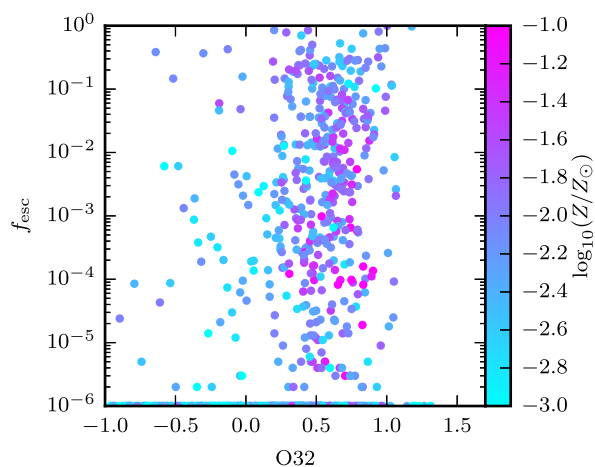
Figure 5. O32 versus R23 for galaxies in our simulation compared to observations. The black and red points represent simulated galaxies with $f_{esc} < 10$ per cent and $f_{esc} \geq 10$ per cent, respectively. The shaded 2D histogram represents local galaxies as observed by SDSS. Blue points represent $z \sim 2$–3 galaxies from Strom et al. (2017), while the cyan points show the $z \sim 3$ LyC leakers from Nakajima et al. (2020). We also show the non-leakers from Nakajima et al. (2020) in magenta. Our simulated galaxies overlap similar regions on this diagram to the $z \sim 3$ galaxies from Nakajima et al. (2020); however, this plane is not useful for classifying LyC leakers from non-leakers.

diagram compared to typical, local star-forming galaxies. Hence, they argue that identifying systems with significant [S II] deficits is a useful tool for selecting those that have an ISM that is optically thin to ionizing radiation.
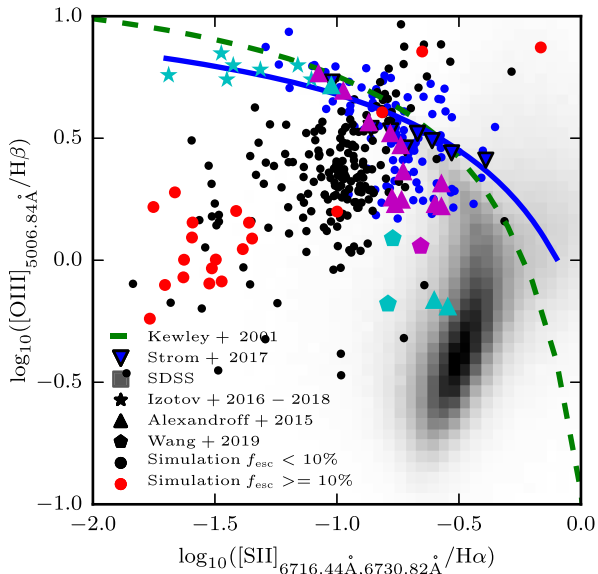
In Fig. 7, we plot the [S II] BPT diagram for galaxies in our simulations compared to observations. Our simulated high-redshift systems (i.e. with $12 \geq z \geq 9.2$) do show an offset on this diagram towards lower [S II]/Hα, more consistent with the expectations of $z \sim 2$–3 galaxies (see Strom et al. 2017; Katz et al. 2019b). The black points on this plot represent simulated systems with $f_{esc} < 10$ per cent, while the red points represent the simulated systems with $f_{esc} \geq 10$ per cent. We find that the majority of leakers have a systematically lower [S II]/Hα than the non-leakers; however, these systems also tend to have systematically lower [O III]/Hβ as well, which was not observed by Wang et al. (2019). This behaviour is indicative of low metallicity (which for our systems ranges from $10^{-4} \leq Z/Z_\odot \leq 0.4$). From Fig. 1, there is a strong trend between f_{esc} and mass such that the lower mass haloes tend to have higher f_{esc}. These lower mass haloes also tend to have lower metallicity and thus the trend that we see in Fig. 7 is reflecting this effect. The reionization-era simulations of Barrow et al. (2017) also show numerous galaxies scattering to the bottom left of the [S II] BPT diagram, consistent with this being a metallicity effect.

The [S II]-deficient LyC leakers presented in Wang et al. (2019), Alexandroff et al. (2015), and Izotov et al. (2016b, 2018a) do not overlap the regions of the [S II] BPT diagram that are populated by our simulated high-redshift LyC leakers and thus we can conclude that these observed, low-redshift [S II]-deficient galaxies may not necessarily be analogues of the galaxies that reionized the Universe. Nevertheless, it is possible that low-redshift galaxies will be found that populate the same regions of the [S II] BPT diagram that are occupied by our simulated-high-redshift galaxies. The simulated

Figure 6. Top: O32 versus f_{esc} for galaxies in our simulation coloured by metallicity. All galaxies with $f_{esc} < 10^{-6}$ have been set to that value so they are visible on the plot. Middle: O32 versus metallicity for galaxies in our simulation coloured by ionizing luminosity. Bottom: O32 versus H I ionization rate in the halo for galaxies in our simulation coloured by escape fraction. While it is difficult to identify trends in the O32–f_{esc} plane, it is clear that the galaxies in our simulation with the highest O32 (e.g. O32 > 1) have lower metallicities and higher H I ionization rates.

Figure 7. [S II] BPT diagram for simulated galaxies versus observations. Local SDSS galaxies are shown as the grey 2D histogram. The dashed green line shows the maximum theoretical starburst line from Kewley et al. (2001). Blue points and the blue line and triangles represent observations of individual star-forming galaxies and the mean relation at $z = 2$–3 from Strom et al. (2017). Individual galaxies from our simulation are shown as circles. Black circles represent systems that have $f_{esc} < 10$ per cent, while red points show galaxies in our simulations that are leaking a considerable amount of LyC radiation and have $f_{esc} \geq 10$ per cent. Cyan stars represent leaky Green Pea galaxies from Izotov et al. (2016a,b, 2018a,b). Cyan and magenta triangles represent leaky and non-leaky Lyman-break analogues from Alexandroff et al. (2015). Finally, cyan and magenta pentagons represent leaky and non-leaky star-forming galaxies from Wang et al. (2019). The regions of the S II BPT diagram populated by the simulated high-redshift leakers are clearly different from the low-redshift galaxies. The regions of this diagram populated by simulated leakers and non-leakers are clearly different indicating that if low-redshift systems are identified that overlap with our simulated high-redshift galaxies, this diagnostic could potentially be used to classify them as leakers.

LyC leakers tend to be offset from the non-leakers in this diagram suggesting that it could still be a useful diagnostic for identifying LyC leakers. The most ideal diagnostics that we derive from our simulations are further discussed in the next section.

4 NEW OBSERVATIONAL DIAGNOSTIC TO SELECT LYC LEAKERS

In this section, we use our simulations to try and develop new methods for selecting LyC leakers from observations. We test these new methods against a sample of dwarf galaxies from the local Universe to identify potential local LyC leakers and a sample from $z > 6$ to identify potential galaxies that actively played a role in reionizing the Universe.

4.1 Multiband models

Because we have nearly complete information for each of our simulated galaxies on both nebular and IR emission line luminosities and LyC escape fractions, we aim to develop a diagnostic that will most efficiently differentiate leakers from non-leakers. To do this, we combine sparse modelling with simple machine learning algorithms to efficiently select the emission line combinations that best correlate with f_{esc}. For this experiment, we use the same sample as above and

only include haloes that are sampled by at least 300 dark matter particles that contain at least 20 star particles.
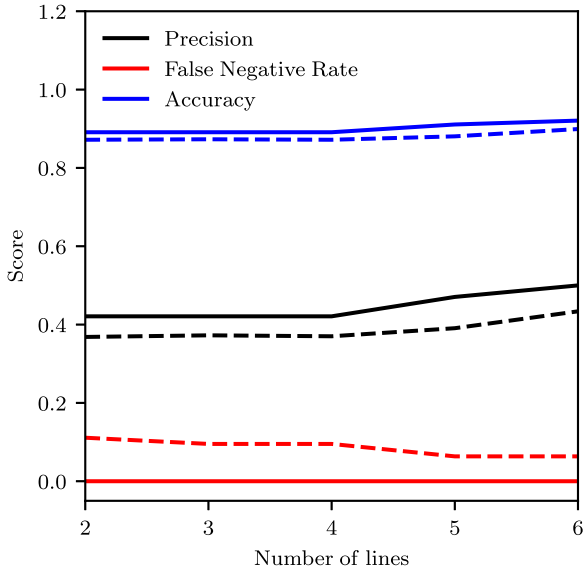
Our method is as follows. We set up a feature matrix containing the logarithm of the emission line luminosities of $[C \, II]_{158 \, \mu m}$, $[O \, III]_{88 \, \mu m}$, $[N \, II]_{6583 \, Å}$, $H\alpha$, $[O \, III]_{5007 \, Å}$, $H\beta$, $[O \, II]_{\lambda\lambda 3727 \, Å}$, $[O \, III]_{4959 \, Å}$, $[S \, II]_{\lambda\lambda 6724 \, Å}$, and $[Ne \, III]_{3869 \, Å}$ for each galaxy in our simulation across all simulation snapshots.[3] The IR lines were selected for this exercise as they represent some of the brightest IR emission lines and are the two commonly observed IR lines in the epoch of reionization. The optical lines were chosen because they appear in common strong line diagnostics. Each galaxy is assigned a status of leaker or non-leaker depending on whether f_{esc} is greater than or less than 10 per cent. We then perform a stratified split of the data into a training set and a test set with a ratio of 80:20 per cent so that each data set has an equal fraction of leakers to non-leakers. The latter set is to determine how well the model generalizes to unseen data.

Once the data have been prepared, we train a logistic regression model to best classify galaxies in the training set as either leakers or non-leakers. This is done by minimizing the binary cross-entropy subject to a weighting scheme where we set class weights proportional to $1 - \frac{N_{class}}{N_{galaxies}}$, where N_{class} is the total number of galaxies in each class. We apply this weighting scheme because there is a high imbalance in the number of leakers to non-leakers in our data set and by weighting our cost function, we can ensure that the scheme is not biased towards predicting more non-leakers correctly.

For practical purposes, we would like to build the best classifier that minimizes the number of lines necessary to make the prediction as the more lines required for the prediction increases the expense of observations. For this reason, we train multiple models where we apply an L1 regularization scheme with different regularization strengths. We use the training set to search a grid of inverse regularization strengths between 0.1 and 10 at intervals of 0.05 and we select the different models such that they have the greatest inverse regularization strength per number of lines. Once we select out the lines that are most informative, we retrain the model without regularization to find the best-fitting model for that set of lines. In principle, adding lines to the model will make it more accurate while decreasing the inverse regularization strength will decrease the accuracy of the model. By sampling multiple models that use different numbers of lines for the prediction, we can balance the cost of the observation with accuracy.

In Fig. 8, we show the accuracy, false negative rate, and precision of the models as a function of the number of lines used in the classifier. For all models, these values are relatively consistent between the training set (dashed lines) and the test set (solid lines) indicating that the models generalize very well. As expected, if we increase the number of lines used for classification, the accuracy (i.e. the fraction of galaxies that we correctly classify as leakers or non-leakers) increases. Likewise, the precision (i.e. the fraction of galaxies that the classifier predicts as leakers that actually are leakers) also increases. When the number of lines used for classification increases from two to six, the precision increases from ∼40 to ∼50 per cent. This value is arguably the most important as it indicates that if we applied our model to real data and then observed the galaxies with a high leaker probability, only 40–50 per cent would turn out to be true leakers.

---

[3] Note that for some galaxies, we predict very low luminosities for the emission lines. This occurs mainly in the haloes with very low mass. In order for this not to impact our training in this work, we set a minimum luminosity of $10 \, L_{\odot}$ for each emission line in this section.
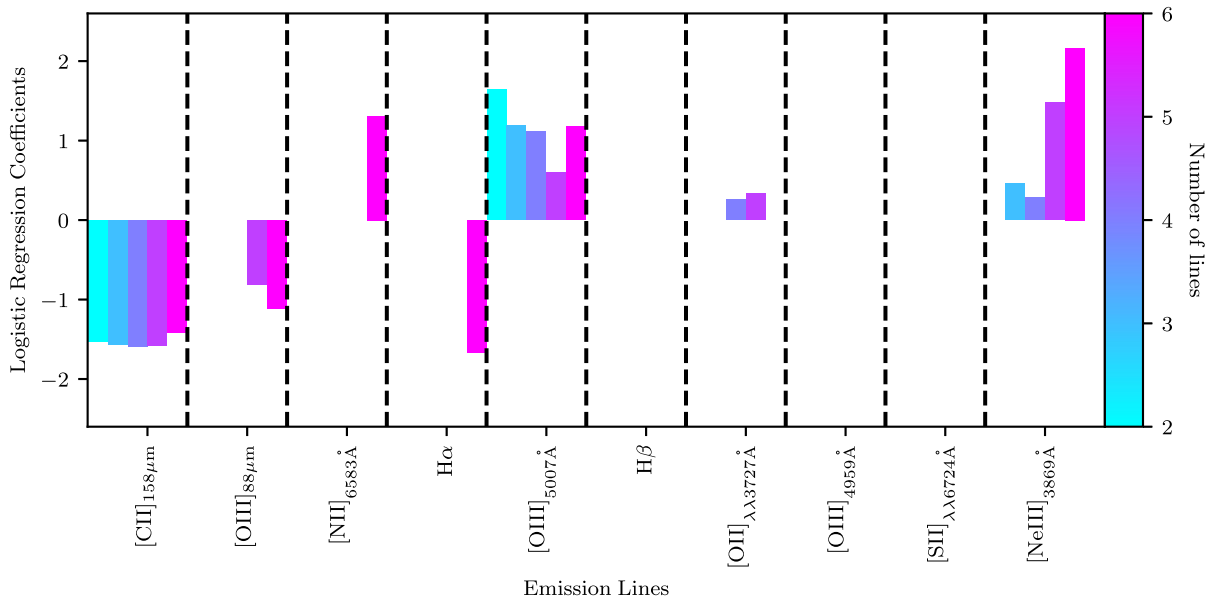
Figure 8. Accuracy (blue), precision (black), and false negative rate (red) as a function of the number of lines in the model for the training set (dashed) and test set (solid). These metrics are fairly consistent between the training and test sets indicating that the models generalize well. Both the precision and accuracy increase with the number of lines used in the model, while the false negative rate remains both low and constant.

Given the difficulty in observing leakers in the local Universe, this high predicted hit rate is very promising. In contrast, the false negative rate (i.e. the fraction of leakers that we misclassify as non-leakers) remains very low ($\lesssim 10$ per cent) for all models. This is due to the fact that we artificially increased the class weights of the leakers when training the model so as to not misclassify them. This certainly impacts the total accuracy of the model and decreases it compared to one where we do not adjust the class weights; however,
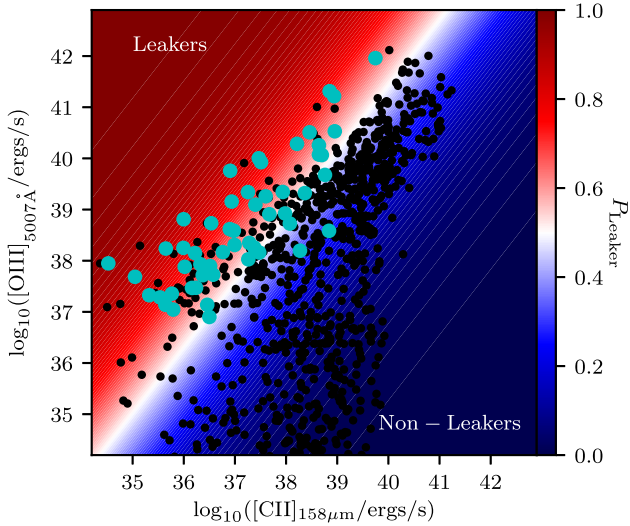
incorrectly classifying true leakers is arguably worse than having a slightly lower precision and accuracy.

By analysing the coefficients in the trained logistic regression model, we can better understand how different emission lines correlate with the LyC escape fraction and which lines can be used to classify leakers and non-leakers. Note that in order for this to work, emission line strengths have to be properly normalized and when training the models, we performed a standard scaling by subtracting the mean and dividing by the standard deviation. In Fig. 9, we show the logistic regression coefficients for each emission line for the different models. The only non-zero coefficients in our two-line model are those of the $[C\,\textsc{ii}]_{158\,\mu m}$ and the $[O\,\textsc{iii}]_{5007\,\text{Å}}$ lines. Because the coefficients are negative and positive, respectively, this indicates that f_{esc} negatively correlates with $[C\,\textsc{ii}]_{158\,\mu m}$ and positively correlated with $[O\,\textsc{iii}]_{5007\,\text{Å}}$. This is expected from our simulations because we know that $[C\,\textsc{ii}]$ emission is highly correlated with the presence of neutral gas (Katz et al. 2017, 2019b), while the higher ionization line $[O\,\textsc{iii}]$ is much more representative of ionized gas and SN feedback (Katz et al. 2019b). However, it is important to note that in the nearby Universe, for higher metallicity systems, a substantial, albeit still subdominant fraction of the $[C\,\textsc{ii}]$ emission is expected to originate from ionized regions (see fig. 9 of Cormier et al. 2019).

In Fig. 10, we show how the two-line model separates the space of $[C\,\textsc{ii}]_{158\,\mu m}$ and $[O\,\textsc{iii}]_{5007\,\text{Å}}$ into regions where it predicts an $f_{esc} \geq 10$ per cent and $f_{esc} < 10$ per cent. The cyan and black points represent leakers and non-leakers in the simulation, respectively, while the red and blue regions show where the model predicts high and low f_{esc}, respectively. The white region represents the dividing line of being classified as a potential leaker by our model. Note that because of the class-weighting scheme we applied, this does not represent an actual probability. The true probability has to be calculated from the results from the test set that show a probability of 40 per cent of being a leaker in the red region (as given by the positive predictive value of the model). This is most easily visualized by looking at the amount of contamination of black points in the red region. Overall, it is clear from Fig. 10 that the logistic regression

Figure 9. Logistic regression coefficients fitted for the different models. Negative values indicate that f_{esc} anticorrelates with the strength of the emission line, while positive values show lines that have a positive correlation with f_{esc}. A zero coefficient represents a line that is not used for the prediction. As the regularization is decreased, the number of lines included in the model increases as given by the colour of the bar.

Figure 10. [C II]$_{158\,\mu m}$ versus [O III]$_{5007\,\text{Å}}$ for our simulated galaxies. The black points represent galaxies with $f_{esc} < 10$ per cent, while the cyan points represent galaxies with $f_{esc} \geq 10$ per cent. The diverging colour map that is underlaid represents the probability that a galaxy is a leaker or a non-leaker from our most simple, two-line logistic regression model. The white band represents the dividing line for when a galaxy is classified as a leaker or a non-leaker in our model. Nearly all of the simulated leakers fall to the left of this line demonstrating that the model has high sensitivity.

model is able to efficiently separate the true leakers from the bulk of the galaxy population.

As the model becomes more complex (due to an increasing inverse regularization strength), other lines are included in the model. After [O III]$_{5007\,\text{Å}}$ and [C II]$_{158\,\mu m}$, the next most important line for classifying a galaxy as a leaker or a non-leaker is the [Ne III]$_{3869\,\text{Å}}$ line. For our most complex models, this line has a coefficient with a similar strength to the [O III]$_{5007\,\text{Å}}$ line. The coefficient for [Ne III]$_{3869\,\text{Å}}$ is also positive indicating that it correlates with a galaxy being a leaker. This is expected because [Ne III]$_{3869\,\text{Å}}$ often traces [O III]$_{5007\,\text{Å}}$ although it has a slightly higher ionization potential and a much higher critical density. Interestingly, as the model becomes more complex, the [O III]$_{88\,\mu m}$ line becomes important but unexpectedly, the coefficient is negative. This line is emitted from gas at much higher densities compared to [O III]$_{5007\,\text{Å}}$ and is perhaps balancing the contribution of the [O III]$_{5007\,\text{Å}}$ line to the prediction by excluding any contribution from the higher density gas. Interestingly, in no model is the [S II]$_{\lambda\lambda 6724\,\text{Å}}$ doublet selected as a good indicator for identifying leakers from non-leakers, nor is [O III]$_{4959\,\text{Å}}$ that contains the same information as [O III]$_{5007\,\text{Å}}$ in terms of it originating from the same regions of the density–temperature phase space. Theoretically, the line ratio of [O III]$_{5007\,\text{Å}}$/[O III]$_{4959\,\text{Å}}$ should be 3 (Storey & Zeippen 2000). What seems to matter most for whether a line is a good indicator of f_{esc} (or whether it provides any additional information compared to other lines) is which regions of the density–temperature phase space that it originates.

To classify an observed galaxy is a leaker or non-leaker, our models can be employed using the following equation:

$$P(f_{esc} \geq 10 \text{ per cent}) = \frac{e^{\Sigma_{i=0}^{10} C_i X_i}}{e^{\Sigma_{i=0}^{10} C_i X_i} + 1}, \tag{1}$$

where the values of the coefficients, C_i, are given for the minimal and complete models in Table 2. X_i represents the $\log_{10}$ of a given line luminosity in solar units normalized by our standard scaling.

The line that corresponds to each coefficient is also given in Table 2. C_0 is not associated with any line as it represents the intercept in the model and thus $X_0 = 1$.

4.2 IR-only model

4.2.1 Model development

While potentially very useful in identifying LyC leakers, one of the main downsides to the above diagnostics is that they use a mix between IR and nebular lines. Because there are no instruments that can simultaneously sample this huge range in frequency, observations from multiple instruments will have to be combined. This in itself can cause issues due to the systematic differences between the various telescopes and there are simply far fewer galaxies that have been observed across such a broad frequency range. For this reason, we develop a second diagnostic that only relies on IR lines; in particular, the [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ emission lines, which have already been suggested to be good indicators of f_{esc} (Inoue et al. 2016). One of the primary advantages of such a model is that the combination of these two lines has already been constrained out to $z = 9.1$ (Laporte et al. 2019), while in the local Universe, larger samples of dwarf galaxies exist for which the luminosity of each of these lines has been measured (e.g. Madden et al. 2013).

The downside of this IR-only model is that it will be less accurate than the two-line model presented above. This is because if it was the optimal model, with the regularization scheme applied above, it would have been naturally selected. Moreover, as described earlier, the [O III]$_{88\,\mu m}$ emission line coefficient from the earlier models was always negative. To first order, one might have assumed this meant that f_{esc} negatively correlates with [O III]$_{88\,\mu m}$, in which case, having a model that combines [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ makes little sense. Rather, our interpretation is that the peak of [O III]$_{88\,\mu m}$ emission occurs at slightly higher densities than [O III]$_{5007\,\text{Å}}$ emission and thus by having a negative coefficient, the model was removing the higher density part of phase space where O III can form. If this interpretation is correct, we should still find success using [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ because [O III]$_{88\,\mu m}$ is sensitive to the feedback in the galaxy, but it will be less optimal than using [O III]$_{5007\,\text{Å}}$ because of the difference in gas density where the emission peaks.

The behaviour of the [O III]$_{88\,\mu m}$ compared to [O III]$_{5007\,\text{Å}}$ may seem curious in our simulation because [O III]$_{5007\,\text{Å}}$ has a considerably higher critical density (7×10^5 cm^{-3}; Baskin & Laor 2005) compared to [O III]$_{88\,\mu m}$ (~ 500 cm^{-3}; Spinoglio et al. 2012). Therefore, one may expect that the line ratio of [O III]$_{5007\,\text{Å}}$ to [O III]$_{88\,\mu m}$ should increase with density (e.g. Keenan & Aggarwal 1990) as collisional deexcitaton should reduce the scaling of the [O III]$_{88\,\mu m}$ line with density above its critical density compared to [O III]$_{5007\,\text{Å}}$. In contrast, our simulations show that the ratio actually increases with density (see fig. 3 of Katz et al. 2019b). This occurs because the regions of our simulation that are bright in both [O III]$_{5007\,\text{Å}}$ and [O III]$_{88\,\mu m}$ rarely reach the critical density of [O III]$_{88\,\mu m}$; hence we do not probe the expected drop off at high densities due to collisional deexcitation. Furthermore, the cosmic microwave background (CMB) at $z = 10$ reduces the emergent line flux of [O III]$_{88\,\mu m}$, particularly at lower densities, while having no impact on [O III]$_{5007\,\text{Å}}$. This explains the behaviour of why [O III]$_{88\,\mu m}$ peaks at higher densities compared to [O III]$_{5007\,\text{Å}}$ in our simulation. We have confirmed that by rerunning the CLOUDY modules with a $z = 0$ CMB background, we recover the expected behaviour.

Fortunately, the best model using only [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ performs only slightly worse than the [C II]$_{158\,\mu m}$ and [O III]$_{5007\,\text{Å}}$

Table 2. Coefficient values and the associated lines for the different models. These should be input into the logistic regression probability equation to classify galaxies as LyC leakers or not.

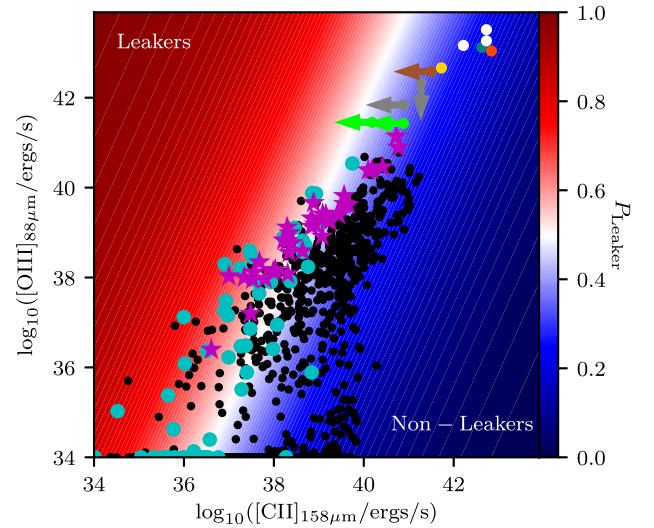| C_i	X_i	μ_i	σ_i	$N_{lines}=2$	$N_{lines}=3$	$N_{lines}=4$	$N_{lines}=5$	$N_{lines}=6$	IR-only model
C_0	1	–	–	-1.27	-1.28	-1.26	-1.33	-1.88	-0.49
C_1	$\frac{\log_{10}([\mathrm{C\,II}]_{158\,\mu m})-\mu_1}{\sigma_1 L_\odot}$	38.24	1.60	-1.53	-1.56	-1.59	-1.57	-1.41	-0.98
C_2	$\frac{\log_{10}([\mathrm{O\,III}]_{88\,\mu m})-\mu_2}{\sigma_2 L_\odot}$	36.14	2.14	0	0	0	-0.81	-1.12	0.59
C_3	$\frac{\log_{10}([\mathrm{N\,II}]_{6583\,\text{Å}})-\mu_3}{\sigma_3 L_\odot}$	36.91	1.84	0	0	0	0	1.31	0
C_4	$\frac{\log_{10}(\mathrm{H\alpha})-\mu_4}{\sigma_4 L_\odot}$	38.79	1.32	0	0	0	0	-1.66	0
C_5	$\frac{\log_{10}([\mathrm{O\,III}]_{5007\,\text{Å}})-\mu_5}{\sigma_5 L_\odot}$	37.47	2.21	1.65	1.19	1.11	0.60	1.18	0
C_6	$\frac{\log_{10}(\mathrm{H\beta})-\mu_6}{\sigma_6 L_\odot}$	38.16	1.31	0	0	0	0	0	0
C_7	$\frac{\log_{10}([\mathrm{O\,II}]_{\lambda\lambda 3727\,\text{Å}})-\mu_7}{\sigma_7 L_\odot}$	37.53	1.75	0	0	0.26	0.34	0	0
C_8	$\frac{\log_{10}([\mathrm{O\,III}]_{4959\,\text{Å}})-\mu_8}{\sigma_8 L_\odot}$	37.03	2.13	0	0	0	0	0	0
C_9	$\frac{\log_{10}([\mathrm{S\,II}]_{\lambda\lambda 6724\,\text{Å}})-\mu_9}{\sigma_9 L_\odot}$	37.46	1.61	0	0	0	0	0	0
C_{10}	$\frac{\log_{10}([\mathrm{Ne\,III}]_{3869\,\text{Å}})-\mu_{10}}{\sigma_{10} L_\odot}$	36.51	1.86	0	0.46	0.29	1.47	2.17	0

model. The parameters for this model can be found in Table 2. We find that the IR-only model has an accuracy of 86 per cent, a precision of 36 per cent and a false negative rate of 6.3 per cent measured on the test set. Both the accuracy and precision have slightly decreased compared to the optimal model, while the false negative rate has increased; however, overall, the performance is not significantly worse. In Fig. 11, we plot $[\mathrm{C\,II}]_{158\,\mu m}$ versus $[\mathrm{O\,III}]_{88\,\mu m}$ for all galaxies in our simulation. LyC leakers that have an $f_{esc} \geq 10$ per cent are shown in cyan, while simulated galaxies with $f_{esc} < 10$ per cent are shown in black. Compared to Fig. 10, one can see that there is less differentiation between leakers and non-leakers in the $[\mathrm{C\,II}]_{158\,\mu m}$–$[\mathrm{O\,III}]_{88\,\mu m}$ space compared to the $[\mathrm{C\,II}]_{158\,\mu m}$–$[\mathrm{O\,III}]_{5007\,\text{Å}}$ space. The underlaid colour map in Fig. 11 represents the probability given by the logistic regression model that an individual system will have an $f_{esc} \geq 10$ per cent. Systems that fall close to the white band have approximately an equal probability of being classified as a LyC leaker. Qualitatively, it is easy to see that the model does a good job in separating leakers from non-leakers.

4.2.2 Comparison with observations

The IR-only model was built in the context of comparing to observations from both low and high redshift. In Fig. 11, we plot 37 galaxies from the local Dwarf Galaxy Survey (DGS; Madden et al. 2013) for which $[\mathrm{C\,II}]_{158\,\mu m}$ and $[\mathrm{O\,III}]_{88\,\mu m}$ have been measured (Cormier et al. 2015) as magenta stars. Interestingly, the local galaxy sample straddles the dividing line between where our model predicts galaxies that have an f_{esc} greater than or less than 10 per cent, allowing us to determine how well it performs on real galaxies.

Unfortunately, observational escape fraction measurements are not available for most DGS galaxies. Thus it is difficult to compare whether the accuracy, precision, and false negative rate of our model coincides with reality. Instead we discuss a few notable galaxies and how our predictions compare with observations.

Of the galaxies predicted to be leakers by our model, NGC 2366 has one of the highest probabilities of having an $f_{esc} \geq 10$ per cent. NGC 2366 is the closest known analogue of a Green Pea galaxy, many of which have already been determined to be LyC leakers (Izotov et al. 2016a,b, 2018a,b). While an escape fraction has not yet been measured for this system, integral field unit (IFU) observations revealed strong ionized outflows that are almost certainly optically thin

Figure 11. $[\mathrm{C\,II}]_{158\,\mu m}$ versus $[\mathrm{O\,III}]_{88\,\mu m}$ for our simulated galaxies. The black points represent galaxies with $f_{esc} < 10$ per cent, while the cyan points represent galaxies with $f_{esc} \geq 10$ per cent. The diverging colour map that is underlaid represents the probability that a galaxy is classified as a leaker or a non-leaker from our IR logistic regression model. The white band represents the dividing line for when a galaxy is classified as a leaker or a non-leaker in our model. Low-redshift observations of local dwarf galaxies are shown as the magenta stars (Madden et al. 2013; Cormier et al. 2015). High-redshift observations of galaxies in the epoch of reionization are shown in lime (Laporte et al. 2019), teal (Hashimoto et al. 2019), grey (Carniani et al. 2017), gold (Bakx et al. 2020), orange-red (companion galaxy only; Walter et al. 2018), brown (Inoue et al. 2016), and white (Harikane et al. 2020). Arrows represent upper limits. Note that for the Carniani et al. (2017) observations, we have treated the [C II] clump separately from the [O III] clump and optical clump on this plot as otherwise, the limits from non-detections would dominate. Our conclusions about this object would not change if we treated both clumps as one system.

and can facilitate the escape of LyC radiation (Micheva et al. 2019). NGC 4861 is another galaxy that our model predicts to be a leaker, although it is less confident about this galaxy compared to NGC 2366. NGC 4861 is similar to NGC 2366 both visually and the fact that it contains an outflow likely driven by SN feedback (van Eymeren

Table 3. Probability that our model classifies a galaxy of having an $f_{esc} \geq 10$ per cent for galaxies from the Dwarf Galaxy Survey (DGS; Madden et al. 2013) as given by our IR-only model.

Galaxy	$P(f_{esc} \geq 10\%)$	Notes
Haro 11	0.22	$f_{esc} = 3.3 \pm 0.7\%$ (Leitet et al. 2011).
Haro 2	0.34	–
Haro 3	0.36	$f_{esc} < 0.1\%$ based on C II (Grimes et al. 2009).
He 2–10	0.34	Confirmed to host an AGN. Observations suggest that the galaxy is optically thick; however, there may be small ionized channels where radiation escapes (Zastrow et al. 2013).
HS 0017+1055	0.57	–
HS 0052+2536	0.27	–
HS 1222+3741	0.54	–
HS 1304+3529	0.40	–
HS 1319+3224	0.61	–
HS 1330+3651	0.39	–
II Zw 40	0.39	Emission is likely to be shock induced and the galaxy is unlikely to be optically thin (Jaskot & Oey 2013).
I Zw 18	0.71	–
Mrk 1089	0.22	–
Mrk 1450	0.61	–
Mrk 153	0.47	–
Mrk 209	0.76	–
Mrk 930	0.29	–
NGC 1140	0.31	$f_{esc} < 1.4\%$ based on C II (Grimes et al. 2009).
NGC 1569	0.45	–
NGC 1705	0.61	Clear evidence that the ISM is being disrupted by stellar feedback; however, detailed observations demonstrate that the outflow is likely optically thick (Zastrow et al. 2013).
NGC 2366	0.72	The closest known analogue of a Green Pea galaxy. IFU observations revealed strong ionized outflows that are almost certainly optically thin and can facilitate the escape of LyC radiation (Micheva et al. 2019).
NGC 4214	0.47	Some of the UV radiation may be escaping based on the dust SED (Hermelo et al. 2013).
NGC 4861	0.63	Similar to NGC 2366 both visually and the fact that it contains an outflow likely driven by SN feedback (van Eymeren et al. 2009).
NGC 5253	0.49	There is an ionization cone that suggests LyC leakage, but it is not clear how much (Zastrow et al. 2011).
NGC 625	0.56	–
Pox 186	0.83	–
SBS 0335−052	0.64	No evidence of emission beyond the Lyman limit (Grimes et al. 2009).
SBS 1159+545	0.60	–
SBS 1415+437	0.64	–
SBS 1533+574	0.42	–
Tol 1214−277	0.51	Lyα observations are highly suggestive that there might be LyC leakage (Verhamme et al. 2015).
UGC 4483	0.76	–
UM 133	0.50	–
UM 311	0.35	–
UM 448	0.19	–
UM 461	0.72	–
VII Zw 403	0.63	–

et al. 2009). Once again, no escape fraction has been measured for this galaxy but it is known from simulations that strong stellar feedback is needed to facilitate the escape of ionizing photons and it is encouraging that our model has selected these two galaxies, both of which show evidence for being disrupted by feedback. NGC 1705 is another example of a galaxy selected to be a leaker by our model. It has a similar probability to that of NGC 4861. Once again, in this system, there is clear evidence that the ISM is being disrupted by stellar feedback; however, detailed observations demonstrate that the outflow is likely optically thick (Zastrow et al. 2013).

One of the more well-studied galaxies in the context of escaping LyC photons is Haro 11. Our model is fairly confident that Haro 11 has $f_{esc} < 10$ per cent, assigning it a 22 per cent probability of having $f_{esc} > 10$ per cent. Interestingly, LyC radiation has been directly observed for this system and the estimate is that $f_{esc} = 3.3$ per cent (Leitet et al. 2011), confirming the prediction from our model. The model similarly predicts minimal LyC leakage in both Haro 2 and also Haro 3.

Finally, our model predicts two other systems as sitting right on the boundary between having $f_{esc} \geq 10$ per cent. These two systems are Tol 1214−277 and NGC 5253. In the case of Tol 1214−277, Lyα observations are highly suggestive that there might be LyC leakage (Verhamme et al. 2015) and for NGC 5253, a dwarf starburst galaxy, there is an ionization cone that suggests LyC leakage as well, but it is not clear how much (Zastrow et al. 2011).

Our model is of course not perfect as given by the fact that the false negative rate is non-zero and the accuracy is not 100 per cent. For example, we predict that SBS 0335−052 has $f_{esc} \geq 10$ per cent, however, the observations from Grimes et al. (2009) show no evidence of emission beyond the Lyman limit. It should be noted that this work also did not infer the detection of LyC emission from Haro 11, which is contradictory to the conclusions from Leitet et al. (2011) and viewing angle will of course play a role. Nevertheless, it must be noted that we predict a contamination rate of up to 65 per cent among the candidates we predict as leakers so some incorrect predictions are expected. A summary

of all systems that we make predictions for can be found in Table 3.

Since the Atacama Large Millimeter/submillimeter Array (ALMA) has come fully online, samples of high-redshift galaxies with observations of both [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ have been made up to $z = 9.1$ (e.g. Laporte et al. 2019). Thus our IR-only model can be used to assess whether or not the galaxies that have been observed directly in the epoch of reionization are potentially the sources of reionization. In Fig. 11, we plot measurements or constraints on [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ for 11 galaxies at $z > 6$ in lime (Laporte et al. 2019), teal (Hashimoto et al. 2019), grey (Carniani et al. 2017), gold (Bakx et al. 2020), orange-red (Walter et al. 2018), and brown (Inoue et al. 2016). Most of these systems have emission line luminosities that are significantly higher than the local dwarfs, which is not surprising because they have been observed at very high redshift. Because these luminosities extend beyond the galaxies in our sample, we have to trust that our model extrapolation suffices in this luminosity regime. Assuming that our model extrapolation holds, the two galaxies with definite [C II]$_{158\,\mu m}$ measurements fall in the blue region of Fig. 11, indicating that they are unlikely to be leaking a significant amount of LyC photons into the IGM. In contrast, those systems for which we have only upper limits [C II]$_{158\,\mu m}$ emission are still strong candidates to be LyC leakers.

In particular, MACS 1149_JD1 from Laporte et al. (2019) sits very close to the boundary. The [C II]$_{158\,\mu m}$ emission line luminosity that is shown represents a 3σ upper limit, indicating that it could be far less luminous. If this is confirmed to be the case, based on our modelling, MACS 1149_JD1 represents a prime candidate for a galaxy that actively contributed to the reionization of the Universe. We note that the actual luminosities for the [C II]$_{158\,\mu m}$ and [O III]$_{88\,\mu m}$ emission lines of MACS 1149_JD1 and A2744_YD4 are formally dependent on the amount of magnification that was assumed to be factors of 10 and 2 for each galaxy, respectively. Because the dividing line between being classified as a leaker and non-leaker in our model is not parallel to the one-to-one line, our conclusions about MACS 1149_JD1 and A2744_YD4 will change if the actual magnification is much different from what is assumed. Furthermore, we note that a recent reanalysis of ALMA data for MACS 1149_JD1 and A2744_YD4 has revealed [C II] detections that are more spatially extended compared to the [O III] (Carniani et al. 2020). These results would place these galaxies much closer the low-redshift [C II]–star formation rate (SFR) relation, consistent with the predictions from our simulations (Katz et al. 2019b); however, it would move the galaxies to the right on the [C II]–[O III] diagram, decreasing their probability of being LyC leakers. In particular, this alternative [C II] measurement for MACS 1149_JD1 is three times higher than the previous upper limit (see table 1 of Carniani et al. 2020).

5 DISCUSSION

5.1 Comparison with previous work

Having developed new diagnostics to identify LyC leakers, it is important to consider how these compare to other methods. Earlier it was determined that the population of simulated galaxies was biased such that the fraction of galaxies that have O32 > 0 that have a $f_{esc} > 10$ per cent was approximately 10 times greater than that with O32 < 0. Indeed, the expected relations between O32 and f_{esc} from Faisst (2016) and Izotov et al. (2018b) rapidly rise beyond this value. These relations contradict those from Bassett et al. (2019) that demonstrate that the rapid rise in f_{esc} at high values of O32 is very sensitive to both the metallicity of the system and the ionization

parameter. Our systems tend to have higher ionization parameters and lower metallicities, consistent with systems closer to $z \sim 3$ that favour a much more gradual increase in f_{esc} as a function of O32 (see the right-hand panel of fig. 10 in Bassett et al. 2019). Nevertheless, if we create a model to separate leakers from non-leakers based on only O32 and use O32 = 0 as the dividing line, only one in 10 systems that we follow up on would turn out to be true leakers. This performs significantly worse than our logistic regression model that has a precision of 40–50 per cent.

One can be more aggressive in their selection methods, for example, Izotov et al. (2016a,b, 2018b) preferentially selected compact, low-metallicity star-forming galaxies that have O32 > 0.69. With their selection criteria, they demonstrated that all six low-redshift galaxies in their sample had leaking LyC radiation. If we apply a cut in O32 at O32 = 0.69, rather than 0, we find that our false negative rate increases dramatically from a few per cent to > 50 per cent because many of the simulated LyC leakers have slightly lower O32. Hence this cut is not satisfying either. However, in combination with another f_{esc} indicator, the false negative rate is expected to drop.

Interestingly, our models put very little weight on the Balmer lines such as Hα and Hβ. This is potentially surprising given the fact that the strength of these lines is expected to be dominated by recombination physics and similarly, other models, such as that discussed in Zackrisson, Inoue & Jensen (2013) partially base their estimate on Hβ. Our systems are not outliers in this regard as fig. 12 of Katz et al. (2019b) demonstrates that our simulated galaxies reasonably reproduce the expected Kennicutt relation between the Balmer lines and SFR. If one knew the intrinsic ionizing photon production rate and the luminosity of Hβ, an estimate for f_{esc} could be made. This method has been used in reverse to estimate the ionizing photon production rate from known leakers (Schaerer et al. 2016). Given this exact coupling, it is perhaps curious as to why our model finds that these lines are not the best indicators for escaping LyC radiation.

We believe that there are two clear distinctions that cause our machine learning models to prefer [O III] and [C II]. First, we have defined our model to classify leakers from non-leakers based on a cut at $f_{esc} = 10$ per cent. This is very different from trying to predict the exact value of f_{esc} from spectral properties. In the latter situation, it is possible that a different method would be necessary that may be more akin to using the UV spectral slope and an equivalent width of a Balmer line (Zackrisson et al. 2013). This method is really only sensitive to $f_{esc} \gtrsim 50$ per cent. In practice, it is still very difficult to determine the global f_{esc} from a galaxy observationally due to the strong variations in f_{esc} depending on sightline. Second, it is well established from simulations that f_{esc} is strongly regulated by feedback (e.g. Kimm & Cen 2014; Kimm et al. 2017; Trebitsch et al. 2017; Rosdahl et al. 2018; Ma et al. 2020). Simulations show that SFR, which correlates very strongly with recombination lines (Kennicutt 1998), is not a good indicator of f_{esc} (e.g. Yoo et al. 2020). The emission lines most sensitive to the feedback, among the ones we have measured, are the [O III] lines. In contrast [C II] follows the neutral gas incredibly well. A decrease in neutral gas favours high f_{esc} as does the presence of feedback and hence, our model selects the two lines that best track these quantities.

5.2 Caveats

While our simulations have been successful in modelling many of the properties of the observed high-redshift galaxy population, there

are numerous caveats that should be kept in mind when interpreting our results.

Like all simulations, the ASPEN simulations have both finite spatial and mass resolution that can impact both the escape fraction and the simulated emission line luminosities. Significantly higher resolution will be needed in the ISM to better resolve the individual low column density channels through which ionizing photons escape (Kakiichi & Gronke 2019; Kimm et al. 2019). Likewise, the resolution of our simulation just barely resolves a multiphase ISM that can significantly impact how emission line luminosities are modelled. For the less resolved galaxies, this can smooth over both the ISM and stellar distribution, both of which will impact our results. A resolution study will be needed to confirm the quantitative results presented here. The simulation depends on a set of subgrid feedback models that were designed so that our galaxy population matches an extrapolation of the high-redshift stellar mass–halo mass relation; however, this choice of feedback significantly impacts the state of the ISM and the LyC escape channels. Furthermore, our simulations are not completely representative of the high-redshift galaxy population because we use a zoom-in simulation of a specific environment around a massive high-redshift galaxy. In order to achieve such high resolution for a massive galaxy system, we used a zoom-in technique rather than simulate the entire volume. While this region contains more than 1000 resolved galaxies, a full-box realization will need to be completed in order to ensure that our results generalize. Extensions to simulations like SPHINX (Rosdahl et al. 2018) or Renaissance (O'Shea et al. 2015) that probe the right mass range, redshift, and ISM physics could be ideal laboratories for similar experiments moving forward.

We reemphasize that comparing to observations is non-trivial both because the LyC escape fractions are only observed along a single line of sight and our treatment of dust is fairly simplistic. In order to confirm or refute our model, larger samples of LyC leakers will be needed so that the comparison can be made probabilistically. Dust will not impact the IR emission line luminosities; however, IR emission line luminosities are sensitive to the exact yields we use for SN in our simulation that were based on solar abundances. It is known from observations that the C/O ratio evolves with metallicity (e.g. Pettini et al. 2008) as the production of C is expected to lag behind that of O. This could impact the galaxies from the DGS that we identify as leakers because it would shift all of the simulated galaxies by some amount to the left on Fig. 11 resulting in fewer predicted leakers. More generally, if the SN that enriched the dwarf galaxies in the local Universe or the high-redshift galaxies have yields that are considerably different to what we have used in the simulations, our diagnostics will have a systematic error. Such an effect is expected to explain systematics shifts in the BPT diagram at $z \sim 2$–3 (e.g. Steidel et al. 2016).

Finally, it is important to note that there are some systematic differences between the simulated galaxies and those observed that impact the comparison. In particular, we have computed our CLOUDY models assuming isotropic radiation from the CMB. To limit the computational demand, we have run the models assuming a CMB radiation field at $z = 10$. The CMB impacts the emergent luminosity of the [C II] 158 μm line and the [O III] 88 μm line, particularly in lower density gas (e.g. da Cunha et al. 2013; Lagache, Cousin & Chatzikos 2018). Our model to identify LyC leakers based on IR lines takes a difference in the log luminosities of these lines so the impact of this is decreased to some extent and similarly, the bulk of the emission arises from the higher density gas that is impacted less; however, the DGS galaxies are not subject to this effect in

the way that our simulated galaxies are, representing a systematic uncertainty.

6 CONCLUSIONS

In this work, we have analysed high-redshift galaxies in the state-of-the-art ASPEN simulations (Katz et al. 2019a,b) to determine whether lower redshift LyC leakers (e.g. de Barros et al. 2016; Vanzella et al. 2016, 2018; Bian et al. 2017; Rutkowski et al. 2017) are 'analogues' of their epoch of reionization counterparts and to develop new emission line diagnostics for identifying LyC leakers. Our main conclusions are as follows.

(i) Similar to observations, galaxies in our simulation that have high LyC $f_{\rm esc}$ are biased towards having high O32. We find that high O32 is likely a necessary but not sufficient condition for high $f_{\rm esc}$ in isolated galaxies (i.e. excluding mergers).

(ii) Our models predict that there likely exist two populations in observations that have high O32 but low $f_{\rm esc}$. The first are those where viewing angle causes a deviation of the line-of-sight $f_{\rm esc}$ from the angle-averaged $f_{\rm esc}$, consistent with the explanation from Nakajima et al. (2020). The second are those where a combination of metallicity and ionization parameter allows for high O32 but low $f_{\rm esc}$, consistent with the models of Bassett et al. (2019).

(iii) We find that the simulated high-redshift galaxies also populate the same regions of the R23–O32 plane as the $z \sim 3$ LyC leakers presented in Nakajima et al. (2020) that tend to have metallicities $\lesssim 0.5 \, Z_\odot$. This combined with evidence in the O32–$f_{\rm esc}$ plane suggests that these highly excited O32 emitters seem to be good analogues of galaxies in the epoch of reionization.

(iv) The LyC-emitting [S II]-deficient galaxies discussed in Wang et al. (2019) may not necessarily be analogues of high-redshift galaxies as strong differences in metallicity causes the simulated systems to populate a different region of the [S II] BPT diagram compared with the [S II]-deficient LyC leakers at low redshift. In general, the leakers in our simulations have [S II] deficits but they also have deficits in [O III] indicating that the dominant effect is metallicity/mass rather than a property of the ISM.

(v) We present multiple logistic regression models to better select LyC leakers from observational data. A minimal model that uses only the [C II]$_{158 \, \mu m}$ and [O III]$_{5007 \, \text{Å}}$ lines and has a true positive rate of 40 per cent, while this value increases to 50 per cent when six lines are included in our most complex models. We intend these models to be used to more efficiently identify LyC leaker candidates at low redshift from existing and future surveys.

(vi) We present an IR-only model that uses the [C II]$_{158 \, \mu m}$ and [O III]$_{88 \, \mu m}$ to classify galaxies that have $f_{\rm esc} \geq 10$ per cent and successfully apply this to local galaxies from the DGS (Madden et al. 2013), as well as various high-redshift galaxies that have been observed deep into the epoch of reionization.

(vii) Based on our results from the IR-only model, MACS 1149_JD1 is the best galaxy candidate that has currently been observed that is likely to be actively contributing to the reionization of the Universe.

In addition to being a powerful probe of ISM properties, emission lines are a useful indicator for the escape of ionizing photons in both low- and high-redshift galaxies. In future work, we will extend the models presented here to include Lyα that has been determined, both observationally and theoretically, to be a good indicator of $f_{\rm esc}$ and look how O32 correlates with $f_{\rm esc}$ on cloud scales not resolved by the current simulations.

DATA AVAILABILITY

The data underlying this paper will be shared on reasonable request to the corresponding author.

REFERENCES

Alexandroff R. M., Heckman T. M., Borthakur S., Overzier R., Leitherer C., 2015, ApJ, 810, 104
Bakx T. J. L. C. et al., 2020, MNRAS, 493, 4294
Barrow K. S. S., Wise J. H., Norman M. L., O'Shea B. W., Xu H., 2017, MNRAS, 469, 4863
Baskin A., Laor A., 2005, MNRAS, 358, 1043
Bassett R. et al., 2019, MNRAS, 483, 5223
Bian F., Fan X., McGreer I., Cai Z., Jiang L., 2017, ApJ, 837, L12
Borthakur S., Heckman T. M., Leitherer C., Overzier R. A., 2014, Science, 346, 216
Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, ApJ, 811, 140
Breiman L., 2001, Machine Learning, 45, 5
Bridge C. R. et al., 2010, ApJ, 720, 465
Buchner J., Bauer F. E., 2017, MNRAS, 465, 4348
Carniani S. et al., 2017, A&A, 605, A42
Carniani S. et al., 2020, preprint (arXiv:2006.09402)
Cen R., Kimm T., 2015, ApJ, 801, L25
Commerçon B., Debout V., Teyssier R., 2014, A&A, 563, A11
Cormier D. et al., 2015, A&A, 578, A53
Cormier D. et al., 2019, A&A, 626, A23
Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53
da Cunha E. et al., 2013, ApJ, 766, 13
de Barros S. et al., 2016, A&A, 585, A51
Dijkstra M., Gronke M., Venkatesan A., 2016, ApJ, 828, 71
Dopita M. A., Krauss L. M., Sutherland R. S., Kobayashi C., Lineweaver C. H., 2011, Ap&SS, 335, 345
Dressler A., Henry A., Martin C. L., Sawicki M., McCarthy P., Villaneuva E., 2015, ApJ, 806, 19
Eldridge J. J., Izzard R. G., Tout C. A., 2008, MNRAS, 384, 1109
Faisst A. L., 2016, ApJ, 829, 99
Fan X., Carilli C. L., Keating B., 2006, ARA&A, 44, 415

Ferland G. J. et al., 2017, Rev. Mex. Astron. Astrofis., 53, 385
Fletcher T. J., Tang M., Robertson B. E., Nakajima K., Ellis R. S., Stark D. P., Inoue A., 2019, ApJ, 878, 87
Geen S., Rosdahl J., Blaizot J., Devriendt J., Slyz A., 2015, MNRAS, 448, 3248
Geurts P., Ernst D., Wehenkel L., 2006, Machine Learning, 63, 3
Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399
Grimes J. P. et al., 2009, ApJS, 181, 272
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Guseva N. G., Papaderos P., Izotov Y. I., Noeske K. G., Fricke K. J., 2004, A&A, 421, 519
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Haiman Z., Loeb A., 1998, ApJ, 503, 505
Harikane Y. et al., 2020, ApJ, 896, 93
Hashimoto T. et al., 2019, PASJ, 71, 71
Hermelo I., Lisenfeld U., Relaño M., Tuffs R. J., Popescu C. C., Groves B., 2013, A&A, 549, A70
Ho T. K., 1995, in Proceedings of the Third International Conference on Document Analysis and Recognition – Vol. 1, ICDAR '95. IEEE Computer Society, Washington, DC, p. 278. Available at http://dl.acm.org/citation.cfm?id = 844379.844681
Hollenbach D., McKee C. F., 1979, ApJS, 41, 555
Inoue A. K. et al., 2016, Science, 352, 1559
Izotov Y. I., Schaerer D., Thuan T. X., Worseck G., Guseva N. G., Orlitová I., Verhamme A., 2016a, MNRAS, 461, 3683
Izotov Y. I., Orlitová I., Schaerer D., Thuan T. X., Verhamme A., Guseva N. G., Worseck G., 2016b, Nature, 529, 178
Izotov Y. I., Schaerer D., Worseck G., Guseva N. G., Thuan T. X., Verhamme A., Orlitová I., Fricke K. J., 2018a, MNRAS, 474, 4514
Izotov Y. I., Worseck G., Schaerer D., Guseva N. G., Thuan T. X., Fricke Verhamme A., Orlitová I., 2018b, MNRAS, 478, 4851
Jaskot A. E., Oey M. S., 2013, ApJ, 766, 91
Kakiichi K., Gronke M., 2019, preprint (arXiv:1905.02480)
Katz H., Ricotti M., 2013, MNRAS, 432, 3250
Katz H., Ricotti M., 2014, MNRAS, 444, 2377
Katz H., Kimm T., Sijacki D., Haehnelt M. G., 2017, MNRAS, 468, 4831
Katz H., Laporte N., Ellis R. S., Devriendt J., Slyz A., 2019a, MNRAS, 484, 4054
Katz H. et al., 2019b, MNRAS, 487, 5902
Keenan F. P., Aggarwal K. M., 1990, ApJ, 350, 262
Kennicutt Robert C. J., 1998, ARA&A, 36, 189
Kewley L. J., Dopita M. A., 2002, ApJS, 142, 35
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121
Kimm T., Cen R., 2014, ApJ, 788, 121
Kimm T., Cen R., Devriendt J., Dubois Y., Slyz A., 2015, MNRAS, 451, 2900
Kimm T., Katz H., Haehnelt M., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 466, 4826
Kimm T., Blaizot J., Garel T., Michel-Dansac L., Katz H., Rosdahl J., Verhamme A., Haehnelt M., 2019, MNRAS, 486, 2215
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kroupa P., 2001, MNRAS, 322, 231
Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, MNRAS, 485, L24
Lagache G., Cousin M., Chatzikos M., 2018, A&A, 609, A130
Laporte N. et al., 2019, MNRAS, 487, L81
Leitet E., Bergvall N., Piskunov N., Andersson B. G., 2011, A&A, 532, A107
Leitet E., Bergvall N., Hayes M., Linné S., Zackrisson E., 2013, A&A, 553, A106
Leitherer C., Hernandez S., Lee J. C., Oey M. S., 2016, ApJ, 823, 64
Levermore C. D., 1984, J. Quant. Spectrosc. Radiat. Transf., 31, 149
Ma X., Quataert E., Wetzel A., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2020, preprint (arXiv:2003.05945)
Madau P., Haardt F., 2015, ApJ, 813, L8
Madau P., Rees M. J., Volonteri M., Haardt F., Oh S. P., 2004, ApJ, 604, 484
Madden S. C. et al., 2013, PASP, 125, 600
Mapelli M., Ferrara A., Pierpaoli E., 2006, MNRAS, 369, 1719
Marchi F. et al., 2017, A&A, 601, A73

Matthee J., Sobral D., Best P., Khostovan A. A., Oteo I., Bouwens R., Röttgering H., 2017, MNRAS, 465, 3637

Meštrić U. et al., 2020, MNRAS, 494, 4986

Micheva G., Christian Herenz E., Roth M. M., Östlin G., Girichidis P., 2019, A&A, 623, A145

Mirabel I. F., Dijkstra M., Laurent P., Loeb A., Pritchard J. R., 2011, A&A, 528, A149

Naidu R. P., Forrest B., Oesch P. A., Tran K.-V. H., Holden B. P., 2018, MNRAS, 478, 791

Naidu R. P., Tacchella S., Mason C. A., Bose S., Oesch P. A., Conroy C., 2020, ApJ, 892, 109

Nakajima K., Ouchi M., 2014, MNRAS, 442, 900

Nakajima K., Ellis R. S., Iwata I., Inoue A. K., Kusakabe H., Ouchi M., Robertson B. E., 2016, ApJ, 831, L9

Nakajima K., Ellis R. S., Robertson B. E., Tang M., Stark D. P., 2020, ApJ, 889, 161

O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, ApJ, 807, L12

Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2015, MNRAS, 451, 2544

Pettini M., Zych B. J., Steidel C. C., Chaffee F. H., 2008, MNRAS, 385, 2011

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

Pontzen A., Roškar R., Stinson G., Woods R., 2013, Astrophysics Source Code Library, record ascl:1305.002

Rémy-Ruyer A. et al., 2014, A&A, 563, A31

Ricotti M., 2002, MNRAS, 336, L33

Ricotti M., Ostriker J. P., 2004, MNRAS, 352, 547

Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19

Roe P. L., 1986, Annu. Rev. Fluid Mech., 18, 337

Rosdahl J., Teyssier R., 2015, MNRAS, 449, 4380

Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, MNRAS, 436, 2188

Rosdahl J. et al., 2018, MNRAS, 479, 994

Rosen A., Bregman J. N., 1995, ApJ, 440, 634

Rutkowski M. J. et al., 2016, ApJ, 819, 81

Rutkowski M. J. et al., 2017, ApJ, 841, L27

Schaerer D., Izotov Y. I., Verhamme A., Orlitová I., Thuan T. X., Worseck G., Guseva N. G., 2016, A&A, 591, L8

Schmidt M., 1959, ApJ, 129, 243

Shapley A. E., Steidel C. C., Strom A. L., Bogosavljević M., Reddy N. A., Siana B., Mostardi R. E., Rudie G. C., 2016, ApJ, 826, L24

Siana B. et al., 2007, ApJ, 668, 62

Siana B. et al., 2010, ApJ, 723, 241

Siana B. et al., 2015, ApJ, 804, 17

Spinoglio L., Dasyra K. M., Franceschini A., Gruppioni C., Valiante E., Isaak K., 2012, ApJ, 745, 171

Stanway E. R., Eldridge J. J., Becker G. D., 2016, MNRAS, 456, 485

Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, ApJ, 826, 159

Steidel C. C., Bogosavljević M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, ApJ, 869, 123

Storey P. J., Zeippen C. J., 2000, MNRAS, 312, 813

Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, ApJ, 836, 164

Teyssier R., 2002, A&A, 385, 337

Toro E. F., 2009, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction, 3rd edn. Springer-Verlag, Berlin

Toro E. F., Spruce M., Speares W., 1994, Shock Waves, 4, 25

Trainor R. F., Steidel C. C., Strom A. L., Rudie G. C., 2015, ApJ, 809, 89

Trebitsch M., Blaizot J., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 470, 224

Trebitsch M. et al., 2020, preprint (arXiv:2002.04045)

van Eymeren J., Marcelin M., Koribalski B. S., Dettmar R. J., Bomans D. J., Gach J. L., Balard P., 2009, A&A, 505, 105

Vanzella E. et al., 2015, A&A, 576, A116

Vanzella E. et al., 2016, ApJ, 825, 41

Vanzella E. et al., 2018, MNRAS, 476, L15

Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, A&A, 578, A7

Verhamme A., Orlitová I., Schaerer D., Izotov Y., Worseck G., Thuan T. X., Guseva N., 2017, A&A, 597, A13

Walter F. et al., 2018, ApJ, 869, L22

Wang B., Heckman T. M., Leitherer C., Alexandroff R., Borthakur S., Overzier R. A., 2019, ApJ, 885, 57

Wise J. H., 2019, preprint (arXiv:1907.06653)

Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560

Xu H., Wise J. H., Norman M. L., Ahn K., O'Shea B. W., 2016, ApJ, 833, 84

Yoo T., Kimm T., Rosdahl J., 2020, preprint (arXiv:2001.05508)

Zackrisson E., Inoue A. K., Jensen H., 2013, ApJ, 777, 39

Zastrow J., Oey M. S., Veilleux S., McDonald M., Martin C. L., 2011, ApJ, 741, L17

Zastrow J., Oey M. S., Veilleux S., McDonald M., 2013, ApJ, 779, 76

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