Lyman continuum leakage versus quenching with the *James Webb Space Telescope*: the spectral signatures of quenched star formation activity in reionization-epoch galaxies

Christian Binggeli,[1][*] Erik Zackrisson,[1] Kristiaan Pelckmans,[2] Rubén Cubo,[2] Hannes Jensen[1] and Ikko Shimizu[3]

[1]*Department of Physics and Astronomy, Uppsala University, Box 515, SE-751 20 Uppsala, Sweden*
[2]*Department of Information Technology, Division of Systems and Control, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden*
[3]*Theoretical Astrophysics, Department of Earth and Space Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan*

ABSTRACT

In this paper, we study the effects of a recent drop in star formation rate (SFR) on the spectra of epoch of reionization (EoR) galaxies, and the resulting degeneracy with the spectral features produced by extreme Lyman continuum leakage. In order to study these effects in the wavelength range relevant for the upcoming *James Webb Space Telescope* (*JWST*), we utilize synthetic spectra of simulated EoR galaxies from cosmological simulations together with synthetic spectra of partially quenched mock galaxies. We find that rapid declines in the SFR of EoR galaxies could seriously affect the applicability of methods that utilize the equivalent width of Balmer lines and the ultraviolet spectral slope to assess the escape fraction of EoR galaxies. In order to determine if the aforementioned degeneracy can be avoided by using the overall shape of the spectrum, we generate mock NIRCam observations and utilize a classification algorithm to identify galaxies that have undergone quenching. We find that while there are problematic cases, *JWST*/NIRCam or NIRSpec should be able to reliably identify galaxies with redshifts $z \sim 7$ that have experienced a significant decrease in the SFR (by a factor of 10–100) in the past 50–100 Myr with a success rate $\gtrsim 85$ per cent. We also find that uncertainties in the dust-reddening effects on EoR galaxies significantly affect the performance of the results of the classification algorithm. We argue that studies that aim to characterize the dust extinction law most representative in the EoR would be extremely useful.

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

1 INTRODUCTION

The epoch of reionization (EoR) represents an extremely important yet poorly understood stage in the evolution of the Universe. During the EoR, the Universe was flooded by energetic hydrogen-ionizing Lyman continuum radiation (LyC), which ionized the neutral intergalactic medium (IGM), but the source of this radiation remains to be determined. Star-forming galaxies have emerged as the main candidate for producing the bulk of the ionizing photons required to sustain reionization, but it is not clear whether the early star-forming galaxies produced and leaked enough such radiation.

In the low- and intermediate-redshift Universe, studying the leakage of LyC photons can be done directly by detecting the escaping LyC photons. There are now several studies that detect leaking LyC with $f_{esc} \sim 0.01$–0.1 (e.g. Bergvall et al. 2006; Leitet et al. 2013;

Borthakur et al. 2014; Izotov et al. 2016b, 2016a) at redshifts $z \lesssim 2$. There are also studies that claim extreme ($f_{esc} > 0.4$) leakage (Shapley et al. 2016; Vanzella et al. 2016; Bian et al. 2017; Izotov et al. 2018). Matthee et al. (2017) also present very high escape fractions ($f_{esc} \approx 0.3$–0.45) for a number of individual sources in a large sample of Lyman-α and H α emitters at $z \sim 2$. When stacking the whole sample they find escape fractions $f_{esc} < 0.1$ for both median and mean stacking. However, detecting the LyC photons directly becomes impossible at $z \gtrsim 4$–5 due to absorption in the increasingly neutral IGM as we approach the EoR. A possible work-around for this is to indirectly determine the LyC escape fraction (f_{esc}) using parts of the spectrum that are unaffected by IGM absorption. A couple of methods for indirectly determining the escape fraction have been suggested. A method using the strengths of absorption lines to probe the line-of-sight covering fraction of neutral hydrogen was proposed by Jones et al. (2013) and Leethochawalit et al. (2016). The method has recently been successfully applied to known

[*] E-mail: Christian.binggeli@physics.uu.se

LyC leakers by Gazagnes et al. (2018) and Chisholm et al. (2018). This method is, however, limited by signal-to-noise requirements at sufficient spectral resolution. It has been suggested that high [O III] / [O II] line ratios could be indicative of Lyman continuum leakage through density-bounded regions in local compact star-forming galaxies (Jaskot & Oey 2013). Low-redshift star-forming galaxies selected for high ratios and compactness have been found to show significant escape fractions (Izotov et al. 2016b, 2016a, 2018), see also Chisholm et al. (2018). Meanwhile, Stasińska et al. (2015) have shown that high [O III] / [O II] ratios are not necessarily an indicator of leaking LyC.

Verhamme et al. (2015) propose that the shape of the Lyman-α spectral line may hold information about the escape fraction, which observations of Lyman-α profiles of local LyC emitters seem to support (Verhamme et al. 2015, 2017). However, as the neutral fraction of the IGM increases with redshift, the shape of the Lyman-α line may change due to Lyman-α transmission effects. Thus, it is uncertain how applicable this method will be to reionization-epoch galaxies (Verhamme et al. 2017).

Another method was suggested by Zackrisson, Inoue & Jensen (2013), and later tested on simulated galaxies from cosmological simulations in Zackrisson et al. (2017). This method exploits the fact that ionizing radiation produced by a young hot population of stars will become reprocessed to longer wavelengths in the neutral hydrogen gas surrounding newly formed stars. Thus, the LyC photons will leave an imprint in the rest-frame non-ionizing UV/optical parts of the spectrum. The authors argue that this could in the future be observed using the *James Webb Space Telescope* (*JWST*). By measuring the equivalent width of the Balmer beta line (H β) while simultaneously measuring the ultraviolet spectral slope, one should be able to identify cases of extreme LyC leakage ($f_{esc} \gtrsim 0.5$) from reionization-epoch galaxies (Zackrisson et al. 2017). In Jensen et al. (2016), this method was further expanded upon using machine learning algorithms in order to increase the number of spectral features used to predict the escape fraction using mock *JWST*/NIRSpec observations. For a discussion on the caveats of this method, see Zackrisson et al. (2017)

Since there are indications that star formation rates (SFRs) in galaxies in the EoR are generally increasing overtime (e.g. Finlator, Oppenheimer & Davé 2011), optical recombination lines of these galaxies are expected to be strong. The method by Zackrisson et al. (2013) utilizes this, and therefore, any weak emission lines in galaxies with blue UV slopes will be interpreted as a telltale sign of Lyman continuum leakage. This could, however, make the method sensitive to rapid changes in star formation. Since the emission in these lines depends on the reprocessing of LyC, and LyC production is dominated by the short-lived O and B stars, a drop in the SFR can lead to weaker lines while not necessarily leading to a significantly redder UV slope. While many cosmological simulations predict star formation histories with fairly modest fluctuations in SFR in the early Universe (Finlator et al. 2013; Gnedin 2014; Gnedin & Kaurov 2014; Shimizu et al. 2014), there are simulations that do feature such variations even for galaxies with stellar masses above $M_\star = 10^8 \, M_\odot$ (Kimm et al. 2015; Ma et al. 2015, 2018). Furthermore, the recent discovery of a $z \approx 9.1$ galaxy that may have undergone a significant drop in the SFR (Hashimoto et al. 2018) raises the question of how common these types of galaxies may be in the early Universe.

If one would catch a galaxy soon after it has undergone a rapid drop in SFR, this object could be incorrectly identified as a galaxy with high-LyC leakage by the method described in Zackrisson et al. (2013, 2017) and Jensen et al. (2016).

We utilize simulated galaxies from the Shimizu et al. (2014) cosmological simulation in combination with mock galaxies that have experienced a recent drop in star formation activity in order to study the effect of quenching on the applicability of using emission lines and the UV slope as a probe of the LyC escape fraction. We assess the possibility of identifying quenched galaxies during the EoR using *JWST*/NIRCam and MIRI photometry. Finally, we utilize linear discriminant analysis (LDA) (Hastie, Tibshirani & Friedman 2009, pp. 106–112) as a classification algorithm in order to assess how well we will be able to distinguish between galaxies that have experienced a recent decline in SFR and star-forming galaxies with high escape fractions. We also assess in which cases the information in the spectral energy distribution is insufficient to realistically distinguish the two types. Note that we do not strictly distinguish between 'attenuation' and 'extinction', but use the terms interchangeably when talking about dust-reddening effects.

2 MODEL

Here, we present the models used in this paper and some important properties of these models. More in-depth properties of the model are discussed in Zackrisson et al. (2017).

2.1 Simulated galaxies and SED modelling

We use the cosmological hydrodynamic simulation by Shimizu et al. (2014), which constitutes a subset of the galaxies used in Zackrisson et al. (2017). The simulation is based on the smoothed particle hydrodynamic code GADGET-3, which is an improved version of the public GADGET-2 code (Springel 2005). In the simulation, 2×640^3 dark matter and gas particles are simulated in a comoving volume of $50 \, h^{-1}$ Mpc cube while considering star formation, supernova feedback, and chemical enrichment following Okamoto, Nemmen & Bower (2008). The simulation does not resolve single stars, but rather collections of stars with masses $\approx 10^6 \, M_\odot$ with a Salpeter IMF (Salpeter 1955), which leads to a minimum stellar mass for the simulated galaxies of around $10^7 \, M_\odot$. We extract the star formation histories, internal metallicity distribution, and the dust content of 406 galaxies with stellar masses $M_\star \geq 5 \times 10^8 \, M_\odot$ at redshift $z \approx 7$ and dust-free magnitudes in the range $M_{1500, AB} \approx -24$ to -20 (corresponding to apparent magnitudes $m_{1500, AB} \approx 23$–27 at $z = 7$) from the simulation.

The information about the star formation history and metallicity is stored in star particles representing a collection of stars with mass of $M_\star \approx 10^6 M_\odot$, while the dust content is calculated for each galaxy as a whole in the simulation and given as a prediction of the extinction at 1500 Å (A_{UV}).

The average stellar metallicity of the galaxies in the simulation ranges from $Z = 7 \times 10^{-4}$ to $Z = 6 \times 10^{-3}$, while single SPH particles can have metallicities $Z < 10^{-5}$. In Fig. 1, we show the star formation histories and total stellar mass as a function of cosmic time for three galaxies from the simulations. On average, the galaxies show increasing SFRs with small variations (factor 2–3) on 10 Myr time-scales once the galaxies reach stellar masses around $10^8 \, M_\odot$, with ~ 5 per cent of galaxies displaying larger variations (factor 5–10). This is shown in Fig. 1, where the black and yellow lines show two typical galaxies, while the red line shows the most extreme galaxy extracted at $z = 7$. This galaxy has a rapid change in SFR (about a factor of 10 drop and shortly thereafter a factor of 20 increase) at around 0.74 Gyr after big bang. This type of galaxy represents an extremely rare type of galaxy in the Shimizu et al. (2014) simulation. However,
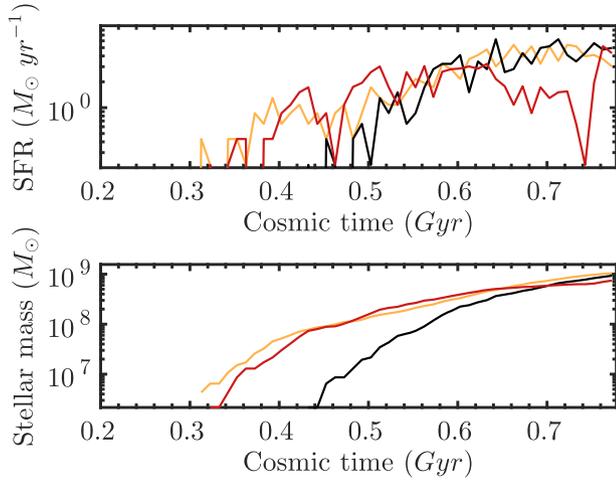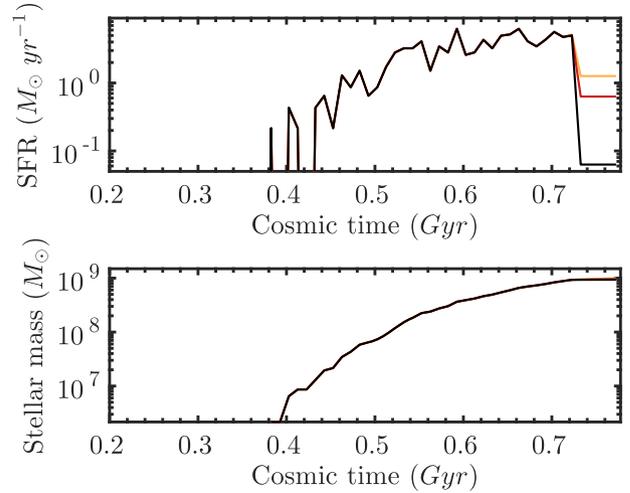
Figure 1. *Top:* Star formation histories binned in 10 Myr bins over cosmic time ($t = 0$ to the age of the Universe at $z = 7$) for two typical (black and yellow) and one extremely stochastically star-forming galaxy *(red)* from the Shimizu et al. (2014) simulations. *Bottom:* Stellar masses for the same two galaxies over cosmic time.

variations of this size and even more extreme cases, with fluctuations in SFR of the order of $\sim 10-100$ over $\sim 10-100$ Myr, are seen more routinely in other simulations (Kimm & Cen 2014; Kimm et al. 2015; Ma et al. 2015, 2018; Trebitsch et al. 2017). Variations of this kind would also be a consequence of a star formation duty cycle of ~ 10 per cent as advocated by Wyithe, Loeb & Oesch (2014).

In this paper, the term '*quenching*' is used to describe a sudden drop in the SFR of galaxies, in contrast to the more regular usage for the term describing a total shut-down in star formation in a galaxy (Harker et al. 2006; Faber et al. 2007). The drop in star formation can be explained for example by the removing of gas due to feedback effects after a recent starburst event (Ma et al. 2018).

In order to generate quenched galaxies, we extract simulated galaxies from the Shimizu et al. (2014) set and lower the star formation with a fixed factor 5,10, and 100 and simulate continuous star formation for 10–100 Myr by aging the already-existing population while adding a less massive population of stars with ages distributed evenly over the time after the quenching takes place and metallicities equal to the average of the simulated stellar population. Thus, the quenched galaxies will have a range of internal metallicity distributions and varying star formation histories up until a significant drop in the SFR occurs. The values for the drop in the SFR and the duration of the quenching were selected to be roughly of the same size as those observed in simulations by Kimm et al. (2015) and Ma et al. (2015, 2018), and should be comparable to the kind of SFR fluctuations seen in simulations by Kimm & Cen (2014) and Trebitsch et al. (2017).

The resulting galaxies are fainter than the simulated ones with dust-free magnitudes in the range $M_{1500,\,AB} \approx -23.5$ to -17.5. In the simulation by Shimizu et al. (2014), this kind of galaxy that is caught in the low SFR stage after quenching is rarer than one in a hundred. In principle, we expect that the effect of this type of quenching on the SED will not be dissimilar to a less massive young starburst on top of a massive underlying aged population, as discussed in Zackrisson et al. (2013). In that study, the authors argue that, for dust-free galaxies, continuum measurements with NIRSpec and MIRI could be used to identify such cases. This could

Figure 2. *Top:* Star formation histories binned in 10 Myr size bins over cosmic time ($t = 0$ to the age of the Universe at $z = 7$) for a mock quenched galaxy shown with a factor of 5 *(yellow)*, 10 *(red)*, and a factor of 100 *(black)* drop in star formation over 50 Myr. *Bottom:* Stellar masses for the same three galaxies over cosmic time. Note that the way the mock galaxies are created they have a mismatch in age with the galaxies from the simulations, which grows with time after the quenching occurs.

however be complicated further when the difference in age between the massive and less massive population is smaller or when taking more complex star formation histories and dust into consideration.

In Fig. 2, we show the star formation history and total stellar mass assembly history for three quenched mock galaxies. We assume a constant SFR once the quenching has occurred.

To get spectral energy distributions (SEDs) from the collection of star particles and the dust content, we use a grid of synthetic spectra for metallicity ranging from $Z = 10^{-7}$ to $Z = 0.03$ generated with the Yggdrasil spectral evolutionary code (Zackrisson et al. 2011). This way, we can assign a spectrum to each star particle and sum over the star particles in the galaxies. To find the closest relevant spectrum for each particle, we perform an interpolation in log(age) and log(Z). The grid of spectra is generated using the BPASS v2.0 binary stellar evolutionary models (Eldridge & Stanway 2009; Stanway, Eldridge & Becker 2016) in combination with spectra from Raiter, Schaerer & Fosbury (2010) for extremely metal poor stars ($Z = 10^{-7}-10^{-5}$). We calculate the Lyman continuum photon production efficiency (the number of ionizing photons produced per unit UV luminosity, ξ_{ion}, in erg^{-1} Hz) for the simulated galaxies. We find values of $\log_{10}(\xi_{ion}/\text{erg}^{-1}\text{ Hz}) \approx 25.2-25.6$ for the individual galaxies, with a mean value for the simulated galaxy set of $\langle \log_{10}(\xi_{ion}/\text{erg}^{-1}\text{ Hz}) \rangle = 25.48$. This result is consistent with several observational studies (e.g. Stark et al. 2015; Bouwens et al. 2015, 2016), motivating the use of the BPASS binary stellar evolutionary models.[1] The nebular emission associated with each star particle is calculated using the Cloudy photoionization code (Ferland et al. 2013) while taking the escape fraction into account and under the assumption that the nebular metallicity is equal to the stellar metallicity. See Zackrisson et al. (2013, 2017) for more in-depth discussion of the procedure. To account for dust, we use the simulated dust amount from the Shimizu et al. (2014) simulations and apply Calzetti et al. (2000) reddening with the same attenuation for the

---

[1]Note that the newer version of the BPASS binary evolution code does predict lower values of ξ_{ion} (Eldridge et al. 2017).

nebular and stellar component ($E(B - V)_{\text{stars}} = E(B - V)_{\text{neb}}$) as the fiducial model. With the amount of dust given in the simulations, this gives an average UV slope $\beta \approx -2.0$ for the whole sample, which is consistent with observations of $z \sim 7$ galaxies (Bouwens et al. 2014). The mean optical extinction in the V-band for these galaxies is $A_V \approx 0.4$ mag with a few galaxies reaching as high as $A_V \approx 1$ mag. In Section 3.2, we discuss the effects of other attenuation and extinction laws, and we there use the standard Calzetti et al. (2000) attenuation law (where $E(B - V)_{\text{stars}} = 0.44 E(B - V)_{\text{neb}}$) and the Pei (1992) SMC extinction law. For the Pei (1992) SMC law, the mean optical extinction in the *V*-band is $A_V \approx 0.2$ mag. While the Calzetti attenuation law gives UV slopes consistent with observations of $z \sim 7$ galaxies, and there are observations of Lyman break galaxies at $z \sim 5$ that point towards a Calzetti-like attenuation law (Koprowski et al. 2018), we note that there are also recent observations that favour a steeper extinction curve for galaxies at $z \sim 7$ (Smit et al. 2018).

In this paper, we assume that all leakage of LyC radiation is occurring through holes in the neutral gas envelope surrounding star-forming regions. This is what is described in Zackrisson et al. (2013) as the ionization-bounded nebula with holes (a.k.a. the picket-fence model). In this type of geometry, the Lyman continuum escape fraction is simply $f_{\text{esc}} = 1 - f_{\text{cov}}$, where f_{cov} is covering fraction of neutral gas. This basically means that the nebular emission will be scaled according to the escape fraction. We calculate spectra for $f_{\text{esc}} = 0, 0.05, 0.10...1.0$. This gives us a total of $\approx 250\,000$ quenched mock galaxies and a total of ≈ 8500 simulated galaxies with varying escape fractions and different star formation histories.

2.2 Mock observations

In order to test whether we can realistically distinguish between quenched and normally star-forming galaxies with variable escape fractions, we create mock *JWST*/NIRCam and MIRI photometry observations of the galaxies and add observational noise to these. For the noise we assume Gaussian noise with a fixed signal-to-noise ratio of S/N = 10 in all filters independent of their sensitivity.

In reality, the S/N will of course depend on the sensitivity in a given filter. For the NIRCam observations, we expect this level of S/N in all filters for around 30 min exposure time for a galaxy with $m_{\text{AB, 1500}} = 27$.[2] For MIRI, we expect a signal-to-noise ratio of at least 10 for galaxies with $m_{\text{AB, 1500}} = 27$ with an exposure of 50 h in the F560W filter. Observations of this type will be performed within the *JWST* guaranteed time observations.[3] In total, we use mock observations in six NIRCam filters (*F115W*, *F150W*, *F200W*, *F227W*, *F356W*, and *F444W*) as well as the *F560W* MIRI filter.

2.3 Linear discriminant analysis (LDA)

The classification algorithm LDA is used in order to classify the galaxies as quenched or as normally star-forming (Hastie et al. 2009, pp. 106–112).

A classification algorithm as LDA aims to recover the probability of an object belonging to a class *C* given input features *X*. LDA is based on Bayes theorem, and makes the parametric assumption that the involved probabilities can be modelled as multivariate Gaussian
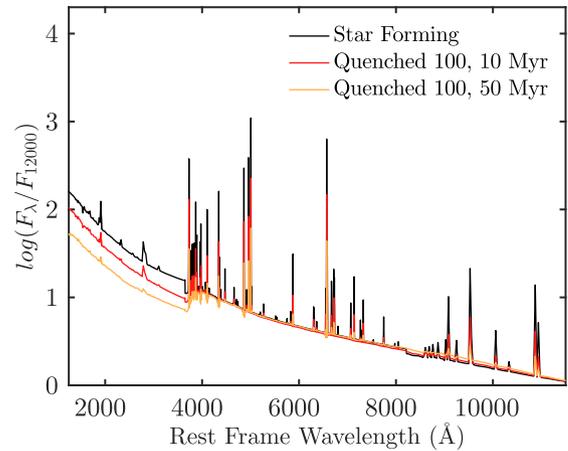
Figure 3. Synthetic spectrum of a single star-forming galaxy ($M_\star \approx 7 \times 10^8$ M$_\odot$) from the simulations *(black)* and the same galaxy after quenching with a factor of 100 has taken place for 10 *(red)* and 50 *(yellow)* Myr. The spectra have been normalized at 12000 Å.

distributions with common covariances but class specific means. Given those quantities, a sample with input features *x* is assigned to class *k* or *l* according to the log-likelihood ratio

$$\log \frac{P(G = k|X = x)}{P(G = l|X = x)} = \log \frac{\pi_k}{\pi_l} - \frac{1}{2}(\mu_k - \mu_l)^T \Sigma^{-1}(\mu_k - \mu_l) + x^T \Sigma^{-1}(\mu_k - \mu_l), \quad (1)$$

where $P(G = k|X = x)$ and $P(G = l|X = x)$ are the probabilities of an observation or object belonging to class *k* and *l*, respectively, given input features *x*. π are the prior probabilities of the classes, μ are the means, and Σ denotes the common covariance matrix. Equation (1) implies that the decision boundary between the classes (where the probabilities are equal) is linear in *x*. Classification to a class *k* and *l* is done according to the maximum conditional probability of belonging to a class.

The training problem amounts to estimating the quantities μ_k, μ_l, and Σ that best describe the classes given some training set of input features. This is performed using maximum likelihood. In this work, the features are NIRCam and MIRI magnitudes and the classes are galaxies that have undergone quenching and those that have not. To avoid skewed data, we undersample the quenched set of galaxies so that we end up with balanced sets of simulated star-forming galaxies and mock quenched galaxies.

3 QUENCHED GALAXIES AND OBSERVABLES

Fig. 3 shows the general effect that quenching has on the SED. As the number of Lyman continuum photons produced drops, the number of such photons that are reprocessed into optical and non-ionizing UV goes down. The emission lines become weaker and the continuum drops in the UV while remaining more or less unchanged at longer wavelengths. The net effect is a reddening of the continuum. 10 Myr after quenching, EW(H β) has dropped a factor of ~ 4 while the UV slope remains more or less the same as in the non-quenched case. 50 Myr after quenching the UV slope has changed significantly while EW(H β) has dropped further. However, the UV slope does not experience extreme reddening due to the quenching. The average reddening of the UV slope over time is only $\Delta\beta \sim 0.2$ and ~ 0.5 for a factor of 10 and 100 drop in SFR over 100 Myr, respectively. While this slight change in the UV

slope could be used to identify cases of quenching if dust effects are ignored, the reddening of the UV slope by dust if one assumes a Calzetti attenuation law for the galaxies is of the same order. The quenched galaxies would thus be indistinguishable from extreme leakers using only EW(H β) and the UV slope β as diagnostics for escape fraction.

Examples of this degeneracy between extreme leakers and quenched galaxies are shown in Fig. 4. This figure displays three pairs of simulated and quenched galaxies, where the galaxies in each pair have basically identical EW(H β) and UV slopes. The pairs in *(a)* and *(b)* in this figure show excess flux in the continuum in *JWST* NIRSpec and NIRCam wavelengths. This is seen especially around 4000 Å. In *(c)*, the galaxies look virtually identical in this wavelength range while the dustier star-forming galaxy is brighter in the rest-frame near-infrared relative to the quenched galaxy. In Section 3.1, we test whether these differences in the continuum are sufficient to identify quenched galaxies when observational noise is taken into account.

3.1 NIRCam + MIRI to identify quenched galaxies

In order to determine whether we can identify if a galaxy has undergone recent quenching or not, we use the calculated magnitudes in NIRCam and MIRI. The distribution of quenched and simulated galaxies in three filters is shown as a colour–colour diagram in Fig. 5. In the noise-free case (panel 1, Fig. 5) the quenched galaxies are largely separated from the normal galaxies. However, when adding noise in this sample, the overlap between the two classes becomes large (panel 2, Fig. 5). In this case we assume that the extinction law is well characterized. We perform 10-fold cross-validation with LDA on the whole set with all the available NIRCam filters and the MIRI F560W filter. This yields a mean accuracy or total accuracy of TA = 81 ± 3 per cent (2σ error), where the accuracy for one validation iteration is defined as

$$A = \frac{TQ + TN}{TQ + TN + FQ + FN}. \tag{2}$$

We define the recall, or true positive rate (TPR), the precision, or positive predictive value (PPV), where we use quenched as positive

$$TPR = \frac{TQ}{TQ + FN}, \tag{3}$$

$$PPV = \frac{TQ}{TQ + FQ}. \tag{4}$$

And the corresponding quantities, specificity, or true negative rate (TNR), and negative predictive value (NPV) for the simulated star forming, or 'normal' galaxies

$$TNR = \frac{TN}{TN + FQ}, \tag{5}$$

$$NPV = \frac{TN}{TN + FN}, \tag{6}$$

where TQ and TN are true quenched and normal, respectively, (those galaxies that are correctly classified as either quenched or normal), and FQ and FN are false quenched and normal, respectively (i.e. those galaxies that are falsely identified as quenched and normal). Table 1 shows the confusion matrix when the model is applied to a test set. This is constructed by randomly taking out 20 per cent of the objects from the training data, meaning that the algorithm has never 'seen' these objects before. In this case, 262 'normal' star-forming galaxies from the simulations are misclassified, while 1439 are

correctly classified. The corresponding numbers for the quenched galaxies are 377 and 1333, which indicates that the algorithm performs slightly better on the normal galaxies with a specificity of approximately 85 per cent and a NPV of 79 per cent, while the recall for the quenched galaxies is 78 per cent and the precision is 84 per cent. This means that about 15 per cent of the normal galaxies will be incorrectly identified as quenched and about 22 per cent of the quenched galaxies will be incorrectly classified as normal. The receiver operating characteristic curve (ROC curve) in Fig. 6 displays how the TPR (recall) and the false positive rate (FPR) changes when changing the threshold used for the classification, together with the area under the curve (AUC). Ideally, one would want zero FPR and maximal TPR, which would mean you have a perfect classifier (i.e. you have no false positive and thus all your positives are true). The dashed line in the ROC curve (Fig. 6) is the line of no discrimination, getting an ROC curve of a classifier along this line would mean our classifier is classifying at random. Having an AUC of unity corresponds to perfect classification, while 0.5 would correspond to a classifier which makes random guesses.

If testing is done only on galaxies that have EW(H β) < 50 Å (which corresponds to galaxies with a 50 per cent escape fraction according to the method of Zackrisson et al. 2013, 2017), we get the confusion matrix shown in Table 2. We see that the specificity (and NPV) in this case is approximately 87 per cent (82 per cent), while the recall (and precision) are 80 per cent (86 per cent).

The general trend of increasing the number of photometric filters used to train and test the algorithm is to increase the total accuracy. However, excluding the MIRI F560W filter from the classification does not significantly affect the result, indicating that it holds very little additional information about the quenching.

In order to identify the most problematic cases of quenching, and for identifying when the algorithm would break down, we create testing sets for each quenching factor and quenching duration. These sets are not used in the respective training phase, hence the algorithm has not 'seen' the objects before. The algorithm is once again trained on a set of mixed quenched and simulated star-forming objects, but now tested onset only consisting of quenched galaxies with certain quenching factor and quenching duration. The results of this test are shown in Fig. 7. Not surprisingly, the algorithm performs better when quenching is stronger and the quenching time-scale is longer, i.e. when the galaxies exhibit clearer signs of aging. If the galaxies are observed only 10 Myr after quenching takes place, the model only achieves a recall of 44–59 per cent depending on the level of quenching, meaning that these galaxies will generally not be distinguishable from simulated star-forming galaxies. Furthermore, it appears that galaxies that have undergone a weak quenching (factor 5 drop in SFR) are problematic. However, for the most extreme cases of quenching (factor 100), a recall of 94 per cent is achieved even after a relatively short quenching duration ($\sim$30 Myr). The recall rate for galaxies that have undergone strong quenching (factor 10–100) after 50 Myr is $\gtrsim$85 per cent.

3.2 Effects of dust handling

If we have access to *JWST*/NIRSpec observations, a correction for dust reddening could in principle be done for galaxies using H β and H γ for $z \sim 6-9$. However, in cases with high escape fractions ($f_{esc} \geq 0.5$), H γ is relatively weak (EW(H γ) < 30 Å) for a majority of the simulated galaxies. Furthermore, as mentioned in Section 2.1, the new version of BPASS binary evolution model predicts lower values of ξ_{ion}, which would lead to weaker emission lines. This and the fact that H α will be only detectable out to $z \sim 7.2$ makes it
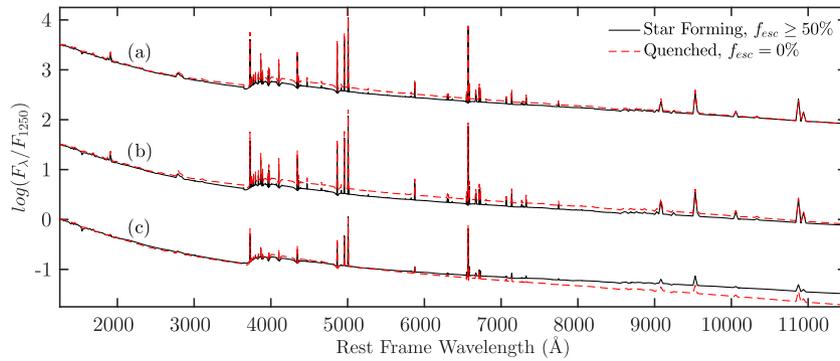
Figure 4. Synthetic spectra showing three examples of the degeneracy between escape fraction and SFR in dusty star-forming galaxies (black) with extreme escape fractions and quenched galaxies (red) with zero escape fractions and lower dust content. *(a)*: Simulated star-forming galaxy with $f_{esc} = 50$ per cent, $A_V \approx 0.41$ mag, and mock galaxy quenched with a factor of 5 with $A_V \approx 0.15$ mag. *(b)*: Simulated star-forming galaxy with $f_{esc} = 70$ per cent, $A_V \approx 0.49$ mag, and mock galaxy quenched with a factor of 10 and $A_V \approx 0.21$ mag. *(c)*: Simulated star-forming galaxy with $f_{esc} = 90$ per cent, $A_V \approx 0.51$ mag, and mock galaxy quenched with a factor of 100 and $A_V \approx 0.09$ mag. All quenched galaxies are observed 40 Myr after the quenching takes place and the Calzetti attenuation law, where $E(B-V)_{stars} = E(B-V)_{neb}$ was used for all galaxies. For clarity, the spectra have been normalized at 1250 Å, *(a)* and *(b)* have been shifted upwards.
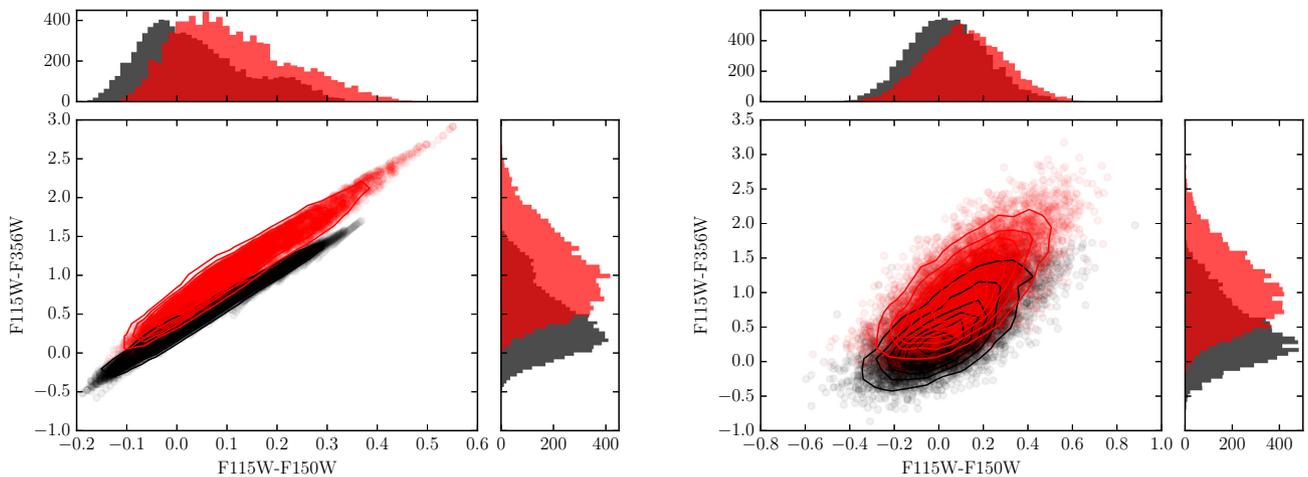

Figure 5. Colour–colour diagrams showing the noise-free (left) distribution and the noisy (SNR = 10) distribution (right) of the quenched galaxies (red) and simulated (star-forming) galaxies (black).

Table 1. Confusion matrix giving the true class *(true)* versus the class that the LDA algorithm predicts *(predicted)* for a testing set consisting of 20 per cent of the total set.

		Predicted	
		Normal	Quenched
True	Normal	1439	262
	Quenched	377	1333

unclear whether or not we can actually reliably correct for dust in these galaxies. We therefore assume that we cannot apply any dust correction.

In Table 3 we show results for the recall, precision, specificity, and NPV, when the algorithm is applied to sets with different attenuation/extinction laws. We see no significant impact on the performance when switching from testing and training on the Calzetti law, where $E(B-V)_{stars} = E(B-V)_{neb}$ to the standard Calzetti law, where $E(B-V)_{stars} = 0.44E(B-V)_{neb}$. Furthermore, there is no significant difference in performance if the model is trained on either Calzetti law, and tested on the other. Meanwhile, the differ-
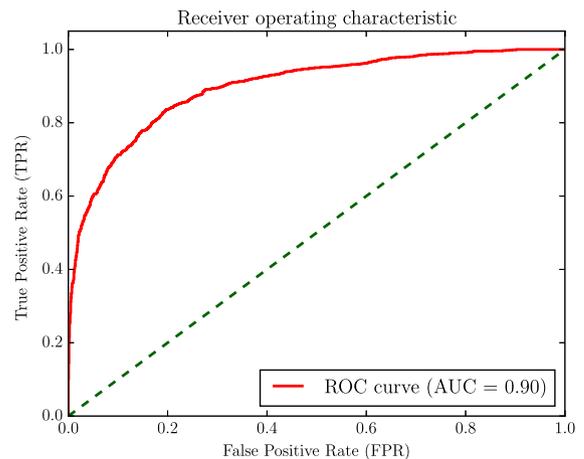
Figure 6. ROC curve for the fiducial set of galaxies. The red line gives the TPR versus the FPR (1−specificity) for the fiducial set as the threshold used for classification is changed. The dashed line displays the line of no discrimination. The area under the ROC curve (AUC) is given in the lower right side of the figure.

Table 2. Confusion matrix showing the true class *(true)* versus the class that the LDA algorithm predicts *(predicted)* when the test set is limited to galaxies that have EW(H β) < 50 Å.

		Predicted	
		Normal	Quenched
True	Normal	429	63
	Quenched	97	395
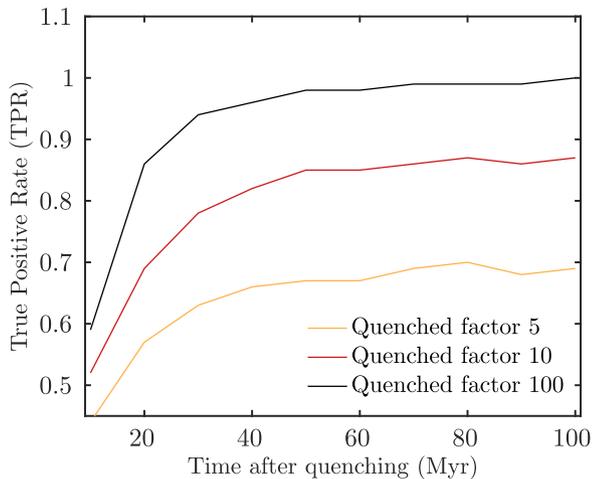
Figure 7. The recall (TPR) of the LDA model as a function of quenching duration for the different factors of quenching. Results shown here are for the fiducial set of objects where no limits have been put on the EW(H β).

Table 3. Recall (TPR), precision (PPV), specificity (TNR), and negative predictive value (NPV) for the LDA model, when it is trained (T) and tested (P) on different dust laws. 'Cal⋆' is the Calzetti law, where $E(B - V)_{stars} = E(B - V)_{neb}$, 'Cal' is the standard Calzetti law and 'SMC' is the Pei (1992) SMC extinction law.

Sets	TPR (%)	PPV (%)	TNR (%)	NPV (%)
T:Cal⋆ P: Cal⋆	78	84	85	79
T:Cal⋆ P: Cal	80	85	86	81
T:Cal⋆ P: SMC	84	87	88	85
T:Cal P: Cal⋆	80	82	83	80
T:Cal P: Cal	80	86	87	81
T:Cal P: SMC	87	83	83	76
T:SMC P: Cal⋆	89	72	66	86
T:SMC P: Cal	90	73	67	87
T:SMC P: SMC	89	89	88	89

ence in performance between training on any of the Calzetti laws and testing on the steeper SMC law is larger, giving a slight increase in performance. We see an overall increase in performance if the model is trained and tested on the SMC law, implying that the steepness of the extinction law leads to less confusion between quenched and 'normal' galaxies. However, we also see a significant decrease in the performance if one assumes an extinction law that is too steep, i.e. if we train on the SMC law and apply the model to a population which is actually experiencing dust-reddening effects by a flatter Calzetti law. In this case, the algorithm will achieve a

TNR of 66–67 per cent, which means that 33–34 per cent of 'normal' star-forming galaxies will be wrongly identified as quenched.

4 DISCUSSION

The previous sections provide evidence for the statement that rapid variations in the SFR in reionziation epoch galaxies will have an impact on the applicability of methods that utilize the emission lines and UV slope to estimate the escape fraction of LyC photons. Quenching of the type discussed in this paper will leave imprints on the SED that are in some degree degenerate with those produced by an extreme escape fraction of LyC photons. However, the effect that quenching has on the overall shape of the SED should, in cases of large decreases in SFR (by a factor of 10–100) in the past 50–100 Myr, allow us to statistically identify quenched galaxies using *JWST* NIRCam and/or NIRSpec observations. For example, consider a situation where our algorithm correctly classifies galaxies in approximately 85 per cent of cases. If we then observe 100 galaxies and find that 15 of those are identified as quenched, it is hard to say anything about how common quenching is. However, if we observe 100 objects and find that 25–35 of those are identified as quenched, it is very likely that there exists a quenched sub-population. Furthermore, by considering a subsample of those galaxies that have H β equivalent width (EW(H β) < 50 Å) we should be able to statistically distinguish between extreme LyC leakage and a recent decrease in SFR. It is still unclear how common this type of quenching is in the early Universe, but if there exists a sub-population of EoR galaxies with star formation histories similar to the one recently claimed by Hashimoto et al. (2018) for a $z \approx 9.1$ object, *JWST* should be able to identify this.

While a sub-population with strong variations in the SFH could be detectable using the *JWST* after roughly 50–100 Myr, distinguishing between cases where weak quenching has taken place (or cases where we catch a galaxy shortly after quenching) and a high escape fraction will be challenging. If we are able to identify a number of strongly quenched galaxies, it is, however, not unlikely that there are galaxies that experience weaker quenching but that are not correctly identified. In principle, this should tell us how applicable the method presented in Zackrisson et al. (2013), Jensen et al. (2016), and Zackrisson et al. (2017) is. If quenching is relatively common, it is likely that there will be cases in which the UV slope and emission lines will lead us to wrongly assume large leakage of LyC photons, and these cases may actually not be distinguishable using only *JWST* observations.

We have shown that the Zackrisson et al. (2013) method is subject to a degeneracy between the LyC escape fraction and star formation activity. Other methods to determine the escape fraction such as high [O III] / [O II] line ratios, Jaskot & Oey (2013), using UV absorption lines to probe the line of sight covering fraction (Jones et al. 2013; Leethochawalit et al. 2016) or constraining the escape fraction using Lyman-α profiles (Verhamme et al. 2015) are not subject to the same degeneracy. These methods do, however, have other limitations that are important to consider when applying the methods to EoR galaxies, as discussed in Section 1.

Meanwhile, identifying a quenched population of galaxies using *JWST* in itself also depends on our understanding of what dust extinction/attenuation law is most representative for the galaxy population at EoR. Even with the JWST/NIRSpec it is not clear if we can reliably perform dust corrections for galaxies at $z \gtrsim 7$. A possible solution that has not been discussed here yet is the combination of MIRI and NIRSpec spectroscopic data. While NIRSpec will not be able to get spectra with H α at redshifts $z \gtrsim 7.2$, MIRI may

be able to provide us with these data. Another possibility could be to characterize dust-reddening effects at lower redshifts, where NIRSpec will have coverage to simultaneously observe H α, H β, and possibly H γ and assume that this dust law is representative even at higher redshifts. In short, any method that can characterize the dust-reddening effects and the variations in the reddening in galaxies during EoR will be extremely useful. If it is the case that there is no extinction law that is representative for the population of galaxies in the early Universe, and in principle every galaxy has a different type of reddening at work, this will strongly limit the applicability of the method that we discuss here. Furthermore, determining which galaxies have undergone quenching also requires that we have a good understanding of how stars are formed and evolve from stellar evolutionary models.

We find that removing the MIRI F560W filter does not significantly affect the performance of the algorithm. The reason for this is most likely that the F560W filter does not probe a significantly different region than the NIRCam F444W filter, and thus provides little additional information about the overall shape of the SED. It is possible that the redder MIRI filters hold information that may improve the identification. However, the limited sensitivity in the F770W and redder filters makes it unclear if these will actually provide information that help the identification of quenched galaxies at these redshifts.

The classification algorithm can also produce results in terms of continuous class memberships (the likelihood for each single galaxy of belonging to a certain class). Computing this for an observed sample of galaxies, one could test how likely we are to find the observed distribution given only a distribution of non-quenched galaxies. This could be used to understand if a quenched sub-population exists in the observed sample of galaxies. However, performing formal significance tests (in the form of P-values) to these statements is challenging as it would involve estimating the underlying distribution of non-quenched galaxies.

We have also tried using a non-linear support vector classifier (Hastie et al. 2009, pp. 417–422) with a radial basis function kernel and tuned parameters, and find no significant difference in performance compared to LDA.

5 SUMMARY AND CONCLUSIONS

(i) We demonstrate that reionization-epoch galaxies that have experienced a recent decrease in star formation activity give rise to *JWST* spectra that are to some extent degenerate to those produced by extreme LyC leakage. This could seriously affect the applicability of the Zackrisson et al. (2013), Jensen et al. (2016), and Zackrisson et al. (2017) method for identifying cases of high-LyC leakage based on low emission line equivalent widths with *JWST*/NIRSpec.

(ii) We show that if we have a good understanding of the general properties of the galaxy population during EoR, we should be able to statistically identify luminous galaxies that have undergone a significant decrease in SFR (by a factor of 10–100) in the past 50–100 Myr using *JWST*/NIRCam and/or NIRSpec observations. Cases with more moderate SFR fluctuations are more difficult to single out, but if there exists a significant population of $z > 6$ galaxies with star formation histories similar to that recently claimed by Hashimoto et al. (2018) for a $z \approx 9.1$ object, then *JWST* should easily be able to pick up on this. Conversely, it should be possible to – in a statistical sense – distinguish galaxies with extreme LyC leakage from objects that have experienced a significant drop in star formation activity.

(iii) While the UV/optical dust extinction is generally assumed to be quite low for $z > 6$ galaxies, the slope of the dust extinction law represents one of the main uncertainties in attempts to break the degeneracy between extreme LyC leakage and quenched star formation. Hence, observations attempting to quantify the details of the dust-reddening effects at $z > 6$ should be considered a high-priority task for *JWST*.

REFERENCES

Bergvall N., Zackrisson E., Andersson B.-G., Arnberg D., Masegosa J., Östlin G., 2006, A&A, 448, 513
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
Bouwens R. J., Smit R., Labbé I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, ApJ, 831, 176
Bouwens R. J. et al., 2014, ApJ, 793, 115
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Chisholm J. et al., 2018, A&A, preprint (arXiv:1803.03655)
Eldridge J. J., Stanway E. R., 2009, MNRAS, 400, 1019
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, PASA, 34, e058
Faber S. M. et al., 2007, ApJ, 665, 265
Ferland G. J. et al., 2013, RMxAA, 49, 137
Finlator K., Muñoz J. A., Oppenheimer B. D., Oh S. P., Özel F., Davé R., 2013, MNRAS, 436, 1818
Finlator K., Oppenheimer B. D., Davé R., 2011, MNRAS, 410, 1703
Gazagnes S., Chisholm J., Schaerer D., Verhamme A., Rigby J. R., Bayliss M., 2018, preprint (arXiv:1802.06378)
Gnedin N. Y., 2014, ApJ, 793, 29
Gnedin N. Y., Kaurov A. A., 2014, ApJ, 793, 30
Harker J. J., Schiavon R. P., Weiner B. J., Faber S. M., 2006, ApJ, 647, L103
Hashimoto T. et al., 2018, Nature, 557, 392
Hastie T., Tibshirani R., Friedman J. H., 2009, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edn. Springer-Veralg, New York
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., 2018, MNRAS, 474, 4514
Jaskot A. E., Oey M. S., 2013, ApJ, 766, 91
Jensen H., Zackrisson E., Pelckmans K., Binggeli C., Ausmees K., Lundholm U., 2016, ApJ, 827, 5
Jones T. A., Ellis R. S., Schenker M. A., Stark D. P., 2013, ApJ, 779, 52
Kimm T., Cen R., 2014, ApJ, 788, 121
Kimm T., Cen R., Devriendt J., Dubois Y., Slyz A., 2015, MNRAS, 451, 2900
Koprowski M. P. et al., 2018, MNRAS, in press
Leethochawalit N., Jones T. A., Ellis R. S., Stark D. P., Zitrin A., 2016, ApJ, 831, 152

Leitet E., Bergvall N., Hayes M., Linné S., Zackrisson E., 2013, A&A, 553, A106

Ma X., Kasen D., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Kereš D., Murray N., 2015, MNRAS, 453, 960

Ma X. et al., 2018, MNRAS, 478, 1694

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

Okamoto T., Nemmen R. S., Bower R. G., 2008, MNRAS, 385, 161

Pei Y. C., 1992, ApJ, 395, 130

Raiter A., Schaerer D., Fosbury R. A. E., 2010, A&A, 523, A64

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

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

Shimizu I., Inoue A. K., Okamoto T., Yoshida N., 2014, MNRAS, 440, 731

Smit R. et al., 2018, Nature, 553, 178

Springel V., 2005, MNRAS, 364, 1105

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

Stark D. P. et al., 2015, MNRAS, 454, 1393

Stasińska G., Izotov Y., Morisset C., Guseva N., 2015, A&A, 576, A83

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

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

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

Wyithe J. S. B., Loeb A., Oesch P. A., 2014, MNRAS, 439, 1326

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

Zackrisson E., Rydberg C.-E., Schaerer D., Östlin G., Tuli M., 2011, ApJ, 740, 13

Zackrisson E. et al., 2017, ApJ, 836, 78

This paper has been typeset from a T_EX/L^AT_EX file prepared by the author.