

# Cosmic metal density evolution in neutral gas: insights from observations and cosmological simulations

Robert M. Yates,<sup>1\*</sup> Céline Péroux,<sup>2,3</sup> Dylan Nelson<sup>4</sup>

<sup>1</sup>*Department of Physics, University of Surrey, Stag Hill, Guildford, GU2 7XH, UK*

<sup>2</sup>*European Southern Observatory, Karl-Schwarzschildstrasse 2, D-85748 Garching bei München, Germany*

<sup>3</sup>*Aix Marseille Université, CNRS, LAM (Laboratoire d’Astrophysique de Marseille) UMR 7326, 13388, Marseille, France*

<sup>4</sup>*Universität Heidelberg, Zentrum für Astronomie, Institut für theoretische Astrophysik, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany*

## ABSTRACT

We contrast the latest observations of the cosmic metal density in neutral gas ( $\rho_{\text{met,neu}}$ ) with three cosmological galaxy evolution simulations: L-GALAXIES 2020, TNG100, and EAGLE. We find that the fraction of total metals that are in neutral gas is  $< 40$  per cent at  $3 \lesssim z \lesssim 5$  in these simulations, whereas observations of damped Lyman- $\alpha$  (DLA) systems suggest  $\gtrsim 85$  per cent. In all three simulations, hot, low-density gas is also a major contributor to the cosmic metal budget, even at high redshift. By considering the evolution in cosmic SFR density ( $\rho_{\text{SFR}}$ ), neutral gas density ( $\rho_{\text{HI}}$ ), and mean gas-phase metallicity ( $[\langle M/H \rangle]_{\text{neu}}$ ), we determine two possible ways in which the absolute  $\rho_{\text{met,neu}}$  observed in DLAs at high redshift can be matched by simulations: (a) the  $\rho_{\text{SFR}}$  at  $z \gtrsim 3$  is greater than inferred from current FUV observations, or (b) current high-redshift DLA metallicity samples have a higher mean host mass than the overall galaxy population. If the first is correct, TNG100 would match the ensemble data best, however there would be an outstanding tension between the currently observed  $\rho_{\text{SFR}}$  and  $\rho_{\text{met,neu}}$ . If the second is correct, L-GALAXIES 2020 would match the ensemble data best, but would require an increase in neutral gas mass inside subhaloes above  $z \sim 2.5$ . If neither is correct, EAGLE would match the ensemble data best, although at the expense of over-estimating  $[\langle M/H \rangle]_{\text{neu}}$ . Modulo details related to numerical resolution and HI mass modelling in simulations, these incompatibilities highlight current tensions between key observed cosmic properties at high redshift.

**Key words:** galaxies: formation – galaxies: evolution – galaxies: abundances – methods: numerical

## 1 INTRODUCTION

The evolution of the cosmic metal density ( $\rho_{\text{met}}$ ) and cosmic star-formation rate density ( $\rho_{\text{SFR}}$ ) are key diagnostics for understanding galaxy evolution. Combined, they encode the relative significance of all the main physical processes driving galaxy evolution, including gas accretion, star formation, stellar evolution, metal enrichment, and outflows driven by supernovae (SNe) and active galactic nuclei (AGN).

Recently, observational studies of the metal content in and around galaxies have suggested that nearly all the metals in the Universe at  $z \gtrsim 2.5$  reside in regions dominated by neutral gas (Péroux & Howk 2020, hereafter PH20). These regions have HI column densities of  $N(\text{HI}) > 2 \times 10^{20} \text{ cm}^{-2}$  and are probed in absorption against bright background sources as damped Lyman- $\alpha$  (DLA) systems. DLAs are expected to trace the cold and warm neutral medium within galaxies (Wolfe et al. 2005; Bird et al. 2014; Rahmati et al. 2015; Peebles et al. 2019), and do not suffer from the selection biases or metallicity calibration issues that hamper emission-based samples. An analysis of this DLA data by PH20 argues that the cosmic, dust-corrected, metal mass density in neutral gas ( $\rho_{\text{met,neu}}$ ) is only a factor of a few

lower than the total expected metal density in the Universe ( $\rho_{\text{met,tot}}$ ) at  $z \sim 2.5$ , and constitutes  $\gtrsim 85$  per cent at  $z \gtrsim 4$ . These estimates are higher than previously predicted from comparisons between DLAs and Lyman-break galaxies (LBGs) (Rafelski et al. 2014).

Such a result is perhaps surprising, given that the presence of strong galactic winds in highly-star-forming galaxies and AGN hosts at high redshift (e.g. Adelberger et al. 2003; Steidel et al. 2010; Bischetti et al. 2019; Sugahara et al. 2019; Ginolfi et al. 2020; Spilker et al. 2020) suggest an efficient removal of metals from galaxies. This material should enrich the circumgalactic medium (CGM) and intracluster medium (ICM) to high levels by  $z \sim 2$  (e.g. Mushotzky & Loewenstein 1997; Tozzi et al. 2003; Balestra et al. 2007; Anderson et al. 2009; Baldi et al. 2012; McDonald et al. 2016), or at least be transferred to a hot, ionized phase, undetected by DLAs. The recent results on  $\rho_{\text{met,neu}}$  seem to be in tension with this scenario.

A precise census of the distribution of metals in the Universe at high redshift is difficult to obtain from current observations, as the full range of dominant contributors to the total metal budget have yet to be comprehensively measured (see PH20). Therefore, in this work, we turn to large-volume cosmological galaxy evolution simulations to assess the apportionment of metals in and around galaxies back to high redshift from a theoretical perspective. This allows us to determine under which conditions  $\rho_{\text{met,neu}}$  dominates the univer-

\* E-mail: r.m.yates@surrey.ac.uk

sal metal budget above  $z \sim 2.5$ , and what the consequences are for galaxy evolution in general. For this analysis, we compare observational data with predictions from the L-GALAXIES 2020 semi-analytic model (Henriques et al. 2020), the EAGLE hydrodynamical simulation (Schaye et al. 2015), and the TNG100 magnetohydrodynamical simulation (Pillepich et al. 2018b; Springel et al. 2018; Nelson et al. 2018a; Naiman et al. 2018; Marinacci et al. 2018).

This paper is organised as follows: In Section 2, we provide an overview of the observational data compiled and properties calculated by PH20 in order to form their analysis of metal densities. In Section 3, we present the four cosmological-scale galaxy evolution models used in this work: L-GALAXIES 2020 ‘default model’ (DM), L-GALAXIES 2020 ‘modified model’ (MM), EAGLE, and TNG100. In Section 4, we discuss our findings, including the evolution of the cosmic metal, HI, and SFR densities, as well as neutral gas metallicities, in simulations and observations. In Section 5, we outline the prospects for future observations at high redshift. Finally, in Section 6 we provide a summary of our conclusions.

## 2 OBSERVATIONAL DATA

The so-called “missing metals problem” was originally posed some 20 years ago by Pettini (1999) and Pagel (1999) as an order of magnitude shortfall in the co-moving density of metals which have been measured compared with those expected to have been produced by the star formation activity seen in galaxies. At the time, when adding up all the metals which had been measured with some degree of confidence, findings showed that they accounted for no more than  $\sim 10$  per cent of what was expected to have been produced and released by  $z = 2.5$ . Although, more recent studies suggest that this discrepancy is substantially reduced (e.g. Renzini 1998; Ferrara et al. 2005; Bouché et al. 2005, 2006, 2007; Shull et al. 2014).

With the goal of reappraising the missing metals problem, PH20 reviewed this census by making an assessment of the cosmic metal mass density continuously with look-back time. Comprehensive estimates for the metal mass present in stars, the intragroup and intracluster medium (IGrM and ICM), highly and partially ionised gas (i.e. Lyman Limit Systems, LLSs, and sub-DLAs), and neutral gas were made. Of particular interest for this study are the measurements made of the metal density in predominantly neutral gas as a function of redshift.

Absorption lines detected against bright background quasars offer the most compelling way to study the distribution, chemical properties, and metal mass budget of the dense gas in and around galaxies. In these quasar absorbers, the minimum gas density which can be detected is set by the brightness of the background source and the detection efficiency is independent of redshift. In emission, there is significant dispersion in the metallicities implied by different strong-line ratios and, more worryingly, between different calibrations of the same ratios (Kewley & Ellison 2008). Absorption techniques, on the other hand, directly count the number of atoms in a given phase of the gas. Several analyses report that abundances from emission and absorption vary by up to  $\sim 0.6$  dex (e.g. Rahmani et al. 2016, although see Section 4.4 below).

When probing regions dominated by neutral gas, the measurements trace only the dominant ionisation states of the metal element and hydrogen. Similarly, some of the metals are locked into dust grains so that observations of gas-phase metals might be inaccurate due to dust depletion. Historically, studies have attempted to avoid the difficulties associated with the dust bias by observing elements known to be only weakly affected by depletion, such as zinc

(e.g. Vladilo et al. 2000). More recently, a multi-element empirical method has been developed to correct homogeneously for dust depletion (Jenkins 2009; De Cia et al. 2016; Jenkins & Wallerstein 2017; De Cia 2018). Such observations of the neutral phase of gas as traced by DLAs is available up to  $z = 5.3$ , indicating only a mild evolution in mean metallicity over nearly a 13 Gyr timespan compared to other phases. After applying dust corrections, it appears observationally that the metal content in neutral gas dominates over more ionised phases at all redshifts. We also note that such multi-element, dust-corrected metallicities are, on average, 0.2–1.0 dex higher than uncorrected metallicity estimates using only Si or Fe lines. This emphasises both the uncertainty in single-ion DLA metallicity studies (especially at low redshift) and mitigates somewhat any differences in the way DLA metallicities are measured in simulations when comparing to data.

In parallel, an assessment of the total metal production by star formation is available from the compilation of Madau & Dickinson (2014). By integrating  $\rho_{\text{SFR}}$  over time, modulo the return fraction and assuming a yield of metal production, one can calculate the amount of metals expected to be produced by stars at any given epoch. A comparison of these results with the sum of the various contributors described above provides two independent estimates of the global metal budget. If the current direct measurements from DLAs of very large metal densities in neutral gas at high redshift are correct, then this would be sufficient to close the budget above  $z \sim 3$ .

## 3 SIMULATIONS

In this section, we present the three cosmological galaxy evolution simulations used in this work. The key parameters and physical processes pertaining to metal evolution are also summarised in Table 1.

### 3.1 L-GALAXIES 2020

L-GALAXIES 2020<sup>1</sup> (Henriques et al. 2020) is a semi-analytic model of galaxy evolution, built to run on the dark matter (DM) subhalo merger trees of N-body DM simulations. It improves on previous versions of the L-GALAXIES model by including prescriptions for H<sub>2</sub> formation (Fu et al. 2010), detailed chemical enrichment from supernovae (SNe) and stellar winds (Yates et al. 2013), and radially-resolved gas and stellar galactic discs (Fu et al. 2013). In this work, L-GALAXIES 2020 is run on the MILLENNIUM-II simulation (Boylan-Kolchin et al. 2009) with box side length  $96.1 h^{-1} \sim 143$  cMpc and DM particle mass  $7.7 \times 10^6 h^{-1} M_{\odot}$ , in order to be comparable to the volumes of the hydrodynamical simulations which we also study (see below).

The key free parameters in L-GALAXIES 2020, such as the efficiency of conversion of H<sub>2</sub> into stars, the efficiency of supermassive black hole (SMBH) feeding, and others, are simultaneously calibrated following an MCMC formalism (Henriques et al. 2009, 2015) in order to match the observed stellar mass functions and red fractions at  $z = 0 - 2$  and the observed HI mass function at  $z = 0$ . Further information on the baryonic physics implemented into L-GALAXIES 2020 can be found in the supplementary material.<sup>2</sup> Following the method of Angulo & White (2010) and Angulo & Hilbert (2015), the L-GALAXIES 2020 cosmology was rescaled to that of the *Planck*

<sup>1</sup> <https://lgalaxiespublicrelease.github.io>

<sup>2</sup> [https://lgalaxiespublicrelease.github.io/Hen20\\_doc.pdf](https://lgalaxiespublicrelease.github.io/Hen20_doc.pdf)

Simulation	Box size (Mpc/h) <sup>3</sup>	$m_{\text{dm}}$ 10 <sup>6</sup> M <sub>⊙</sub> /h	$m_{\text{b}}$ 10 <sup>6</sup> M <sub>⊙</sub> /h	$h$ km/s/Mpc	Metal yields			SFR	SN feedback
					AGB	SNe-Ia	SNe-II	formula	formula
L-GALAXIES 2020	96.1	7.7	–	0.673	M01	T03	P98	$\dot{\Sigma}_* = \alpha \Sigma_{\text{H}_2} / t_{\text{dyn}}$	$\dot{E}_{\text{SN}} = \eta_{\text{SN}} \epsilon_{\text{SN}} \dot{M}_{\text{R}}$
EAGLE	67.8	6.6	1.2	0.677	M01	T03	P98	$\dot{\Sigma}_* = A \Sigma_{\text{gas}}^n$	$\dot{E}_{\text{SN}} = \chi \epsilon_{\text{SN}} \dot{M}_*$
TNG100	75.0	5.1	0.9	0.677	K10/D14/F14	N97	P98/K06	$\dot{\rho}_* = \rho_{\text{c}} / t_*$	$\dot{E}_{\text{SN}} = \epsilon_{\text{w}} \dot{M}_*$

**Table 1.** The key parameters and equations defining the three simulations considered in this work. More detail is provided in the text in Section 3. The metal yield tables listed refer to the following works: Marigo (2001, M01); Thielemann et al. (2003, T03); Portinari et al. (1998, P98); Karakas (2010, K10); Doherty et al. (2014, D14); Fishlock et al. (2014, F14); Nomoto et al. (1997, N97); Kobayashi et al. (2006, K06). For L-GALAXIES 2020:  $\alpha = 0.06$  is the H<sub>2</sub>-to-stars conversion efficiency,  $\Sigma_{\text{H}_2}$  is the H<sub>2</sub> surface mass density,  $t_{\text{dyn}} = R_{\text{cold}}/V_{\text{max}}$  is the gas disc dynamical time,  $\eta_{\text{SN}} = 0.0149 \text{ M}_{\odot}^{-1}$  is the number of SNe per mass,  $\epsilon_{\text{SN}} = 10^{51} \text{ erg}$  is the energy released per SN, and  $\dot{M}_{\text{R}}$  is the mass return rate by stellar winds and SNe (see Henriques et al. 2020). For EAGLE:  $A = 1.515 \times 10^{-4} \text{ M}_{\odot}/\text{yr}/\text{kpc}^2$  is the gas-to-stars conversion efficiency,  $n = 1.4$  is the slope of the Kennicutt-Schmidt law (Kennicutt 1998),  $\Sigma_{\text{gas}}$  is the surface mass density for gas above a given metallicity-dependent density threshold (Schaye et al. 2010),  $\chi$  is the fraction of SN energy per mass used to drive galactic winds (Springel & Hernquist 2003),  $\epsilon_{\text{SN}} = 10^{51} \text{ erg}$  is the energy released per SN, and  $\dot{M}_*$  is the star formation rate (see Schaye et al. 2015). For TNG100:  $\rho_{\text{c}}$  is the cold gas density above a threshold value,  $t_*$  is the (density-dependent) star formation timescale, which is taken to be proportional to the local dynamical time of the gas, and  $\epsilon_{\text{w}}$  is the specific energy available for wind generation, which is tied to the SN-II energy per mass of stars formed (Pillepich et al. 2018a).

2013 survey, as described by Planck Collaboration et al. (2014b):  $\Omega_{\Lambda,0} = 0.685$ ,  $\Omega_{\text{m},0} = 0.315$ ,  $\Omega_{\text{b},0} = 0.0487$ ,  $\sigma_8 = 0.826$ ,  $n_{\text{s}} = 0.96$ ,  $h = 0.673$ .

One of the main advantages of semi-analytic models is that the prescriptions governing baryonic physics can be easily adapted. Therefore, in this work, we focus on two versions of the L-GALAXIES 2020 model – the ‘default model (DM)’ and the ‘modified model (MM)’. The DM was first presented in Henriques et al. (2020) and assumes 70 percent of the metal released by SNe is immediately mixed with the local ISM, before some is driven out of galaxies through SN-driven galactic outflows. This equates to direct CGM enrichment efficiencies for SNe and AGB winds of  $f_{\text{SNII,hot}} = 0.3$ ,  $f_{\text{SNIa,hot}} = 0.3$ , and  $f_{\text{AGB,hot}} = 0.0$ . The MM was first presented in Yates et al. (2021) and assumes up to 90 per cent of the metal released by SNe is allowed to directly enrich the CGM surrounding galaxies, without first mixing with the ISM, i.e.  $f_{\text{SNII,hot}} = 0.9$ ,  $f_{\text{SNIa,hot}} = 0.8$ , and  $f_{\text{AGB,hot}} = 0.25$ . These two versions of L-GALAXIES 2020 therefore nicely bracket the likely range of direct CGM enrichment efficiencies present in real galaxies.

The stellar yield tables used in both versions of L-GALAXIES 2020 are given in Table 1. These are the same as those used in the EAGLE simulation.

### 3.2 EAGLE

The EAGLE simulation project<sup>3</sup> (Crain et al. 2015; Schaye et al. 2015) is a series of cosmological, hydrodynamical simulations, all run with a modified version of the GADGET-3 code (Springel et al. 2005). Here we exclusively use the largest volume ‘Ref-L100N1504’ simulation, which has a  $67.8 h^{-1} \sim 100 \text{ cMpc}$  side-length box containing  $2 \times 1504^3$  dark matter plus baryonic particles. As a result, the (initial) particle masses are  $1.2 \times 10^6 h^{-1} \text{ M}_{\odot}$  and  $6.6 \times 10^6 h^{-1} \text{ M}_{\odot}$ , for gas and DM, respectively, while the collisionless gravitational softening length is fixed at 0.7 physical kpc at late times ( $z < 2.8$ ). The EAGLE cosmology is based on that from the Planck 2013 survey, as described by Planck Collaboration et al. (2014a):  $\Omega_{\Lambda,0} = 0.693$ ,  $\Omega_{\text{m},0} = 0.307$ ,  $\Omega_{\text{b},0} = 0.04825$ ,  $\sigma_8 = 0.829$ ,  $n_{\text{s}} = 0.9611$ ,  $h = 0.6777$ .

The EAGLE model for galaxy formation physics includes a comprehensive description of the most important processes govern-

ing galaxy growth and evolution. This incorporates star formation, whereby gas above a (metallicity-dependent) threshold density is stochastically converted into stars; stellar evolution tracks the abundances of eleven individual elements, following massive and intermediate-mass stars; radiative cooling and heating are based on these individual elemental abundances, assuming photoionization equilibrium with an optically thin ionizing background radiation field (Wiersma et al. 2009); supermassive black holes are seeded in all sufficiently massive halos, and subsequently grow by accreting gas from their surroundings; stellar feedback as well as supermassive black hole feedback are both treated via a bursty thermal energy release (Dalla Vecchia & Schaye 2012).

### 3.3 TNG100

The IllustrisTNG project<sup>4</sup> (Naiman et al. 2018; Pillepich et al. 2018b; Nelson et al. 2018a; Marinacci et al. 2018; Springel et al. 2018) is a series of large-volume cosmological gravo-magnetohydrodynamics (MHD) simulations, run with a comprehensive model for the physics driving galaxy formation (Weinberger et al. 2017; Pillepich et al. 2018a). The TNG project encompasses three uniform volume boxes: TNG50, TNG100, and TNG300, all run with the exact same physical model, albeit at different numerical resolutions. Here we exclusively use the intermediate TNG100 run, as a compromise between the high-resolution TNG50 realization (Pillepich et al. 2019; Nelson et al. 2019b) and the large-volume TNG300 run. In particular, TNG100 includes  $1820^3$  dark matter particles, plus a roughly equal number of baryonic mass elements spanning gas, stars, and supermassive black holes, within a  $75 h^{-1} \sim 110 \text{ cMpc}$  side-length box.

TNG100 has a baryon mass resolution of  $9.4 \times 10^5 h^{-1} \text{ M}_{\odot}$ , a DM mass resolution of  $5.1 \times 10^6 h^{-1} \text{ M}_{\odot}$ , a collisionless softening of 0.74 physical kpc at  $z = 0$ , and a minimum gas softening of 184 comoving parsecs. Its cosmology is taken from Planck 2015 data (Planck Collaboration 2016), adopting  $\Omega_{\Lambda,0} = 0.6911$ ,  $\Omega_{\text{m},0} = 0.3089$ ,  $\Omega_{\text{b},0} = 0.0486$ ,  $\sigma_8 = 0.8159$ ,  $n_{\text{s}} = 0.9667$  and  $h = 0.6774$ . Halos and subhalos (galaxies) are identified with the SUBFIND algorithm (Springel et al. 2001), as in EAGLE.

TNG uses the AREPO code (Springel 2010) and includes a model for the key processes which regulate the formation of galaxies. Most importantly, feedback from both supernovae and supermassive black

<sup>3</sup> [eagle.strw.leidenuniv.nl](http://eagle.strw.leidenuniv.nl)

<sup>4</sup> [www.tng-project.org](http://www.tng-project.org)

holes (see [Weinberger et al. 2017](#); [Pillepich et al. 2018a](#), for details). The production of metals and the assumed stellar nucleosynthetic yields are taken from a combination of multiple stellar evolution codes, as shown in Table 1, noting that these yield tables differ from those in the original Illustris simulation (see [Pillepich et al. 2018a](#)).

### 3.4 Differences between the simulations

In addition to differing choices in the stellar yield sets used, the three simulations considered here also differ in the number of chemical elements tracked. Both L-GALAXIES 2020 and EAGLE track 11 chemical elements: H, He, C, N, O, Ne, Mg, Si, S, Ca, Fe. The last nine of these (i.e. excluding H and He) constitute  $\sim 98$  per cent of the total metal mass in the photosphere of the Sun ([Asplund et al. 2009](#)), indicating that their sum provides a good approximation to the total metal mass. TNG100 only explicitly tracks 9 chemical elements: H, He, C, N, O, Ne, Mg, Si, Fe, the last seven of which constitute  $\sim 95$  per cent of the total metal mass in the Sun’s photosphere. However, this does not make a significant difference to abundances in TNG100, as the SN-II yields (which dominate the production of the most abundant metals) are also renormalised to match the total mass ejected from the [Portinari et al. \(1998\)](#) tables used in L-GALAXIES 2020 and EAGLE (see [Pillepich et al. 2018a](#)).

The SN feedback models in L-GALAXIES 2020, EAGLE, and TNG100 also differ in a number of substantial ways. This impacts the metal enrichment of gas in and around galaxies. The efficiency with which SN feedback drives galactic outflows can be quantified by the mass loading factor,  $\eta = \dot{M}_w / \dot{M}_*$ , where  $\dot{M}_w$  is the mass outflow rate and  $\dot{M}_*$  is the star formation rate. This factor essentially represents the efficiency of gas expulsion for a given amount of available SN energy (see Table 1).

In EAGLE,  $\eta$  decreases with cosmic time at fixed  $\dot{M}_*$  ([Mitchell et al. 2020](#)), due to the fraction of SN energy made available for outflows being inversely proportional to metallicity ([Furlong et al. 2015](#)). This means that metals can be more efficiently accumulated in the ISM at later times, playing an important role in the rate of metal enrichment, as shown in Section 4.4.

In TNG100,  $\eta$  is typically higher than in EAGLE ([Nelson et al. 2019b](#)), indicating lower metal retention. However, an increased  $\rho_{\text{SFR}}$  can outweigh this effect, through the greater production of metals overall, as discussed in Section 4.2. Also, the redshift evolution of  $\eta$  in TNG100 is somewhat smoother over time [Nelson et al. \(2019b\)](#), at least for galaxies with  $\log_{10}(M_*/M_\odot) < 10.5$  where AGN feedback does not play a significant role ([Ayromlou et al. 2021](#)).

In L-GALAXIES 2020,  $\eta$  values are similar to those in EAGLE at  $z = 0$  ([Yates et al. 2021](#)). However, like TNG100, they are not directly dependent on metallicity, allowing for a more efficient retention of metals in the ISM at high redshift than in EAGLE. For the DM version of L-GALAXIES 2020, this is enough to return high ISM metallicities in low-mass galaxies at early times. For L-GALAXIES 2020 MM however, this effect is strongly offset by the increased metal ejection efficiency (see Section 3.1), which greatly reduces ISM metallicities at high redshift, despite this version of the simulation having lower overall  $\eta$  ([Yates et al. 2021](#)).

The star formation models in the three simulations are more similar to each other (see Table 1). L-GALAXIES 2020 adopts an  $\text{H}_2$ -dependent SFR law, made possible because the partitioning of ISM gas into  $\text{H}$  and  $\text{H}_2$  is modelled on the fly in the simulation ([Fu et al. 2010](#)). EAGLE and TNG100 instead follow the prescription developed by [Springel & Hernquist \(2003\)](#), which relates the star formation rate density to the total density of gas above a certain density threshold of  $n_{\text{H}} \simeq 0.1 \text{cm}^{-3}$  ([Vogelsberger et al. 2013](#); [Schaye et al. 2015](#)).

Both of these approaches reproduce the observed Kennicutt-Schmidt relation ([Kennicutt 1998](#)) for star-forming galaxies.

Finally, the small differences in cosmological parameters assumed between the simulations should only have a negligible impact on their relative results ([Angulo & White 2010](#); [Guo et al. 2013](#); [Henriques et al. 2015](#)), especially when compared to differences in the implementation of baryonic physics.

## 4 RESULTS

### 4.1 Metal density evolution

In this section, we present the main focus of this work – a comparison between the cosmic metal density evolution from observations and the L-GALAXIES 2020, EAGLE, and TNG100 simulations.

#### 4.1.1 Measuring densities

For the simulations, metal densities are calculated at any given redshift as the total metal mass in that particular phase divided by the total comoving volume of the simulation box,

$$\rho_{\text{met,theo}} = \frac{\sum M_{\text{metals}}}{V_{\text{tot}}} . \quad (1)$$

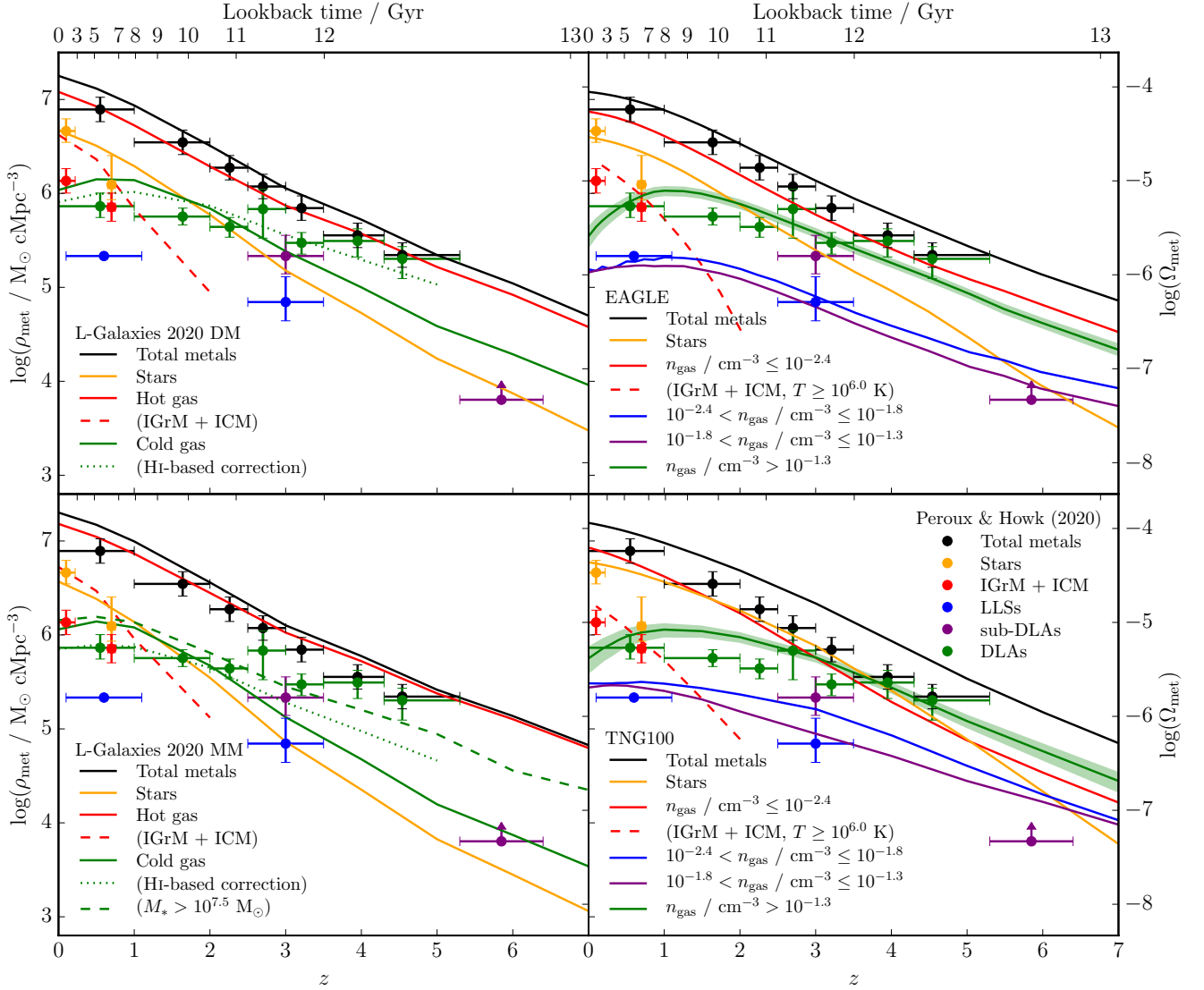
Observationally, neither  $\sum M_{\text{metals}}$  nor  $V_{\text{tot}}$  are readily measurable, so subtly different approaches are taken to infer metal densities. When estimating the total metal density in neutral gas, the following is used

$$\rho_{\text{met,neu,obs}} = \rho_{\text{HI+He}} \cdot \langle M/H \rangle_{\text{DLA}} , \quad (2)$$

where  $\rho_{\text{HI+He}}$  is the universal baryon density of  $\text{H}$  plus helium, as probed by statistical  $\text{H}$  surveys at various redshifts (see [PH20](#) and Section 4.3), and  $\langle M/H \rangle_{\text{DLA}}$  is the  $\text{H}$ -weighted mean metal-to-hydrogen mass ratio from DLAs (see [PH20](#), section 3.4). Eqns. 1 and 2 return essentially the same result (modulo small differences between the mass of hydrogen,  $\text{H}$ + $\text{He}$ , and total neutral gas) in the limit that DLAs are unbiased tracers of neutral gas in the Universe. However, if this is not the case, these two calculations can differ, as discussed in Section 4.4. Previous works have argued that, when selecting DLAs from simulations along random sightlines, and modeling  $\text{H}$  disc sizes and densities, the cosmic neutral gas density inferred is similar to that obtained by considering all  $\text{H}$  in the simulation box ([Berry et al. 2016](#); [Di Gioia et al. 2020](#)).

We highlight here that the cosmic mass densities given by equations Eqns. 1 and 2 are essentially ‘normalised’ total metal masses. This is because the volume considered is always the total comoving volume of the Universe (explicitly in the case of Eqn. 1, and implicitly in the case of Eqn. 2). The actual volume filling fraction of any particular phase does not enter. The equivalent, dimensionless quantity,  $\Omega = \rho / \rho_{\text{crit},0}$ , is therefore sometimes preferred in the literature, where  $\rho_{\text{crit},0} = 3\text{H}_0^2 / 8\pi G = 1.36 \times 10^{11} \text{M}_\odot \text{cMpc}^{-3}$  is the critical mass density of the Universe at  $z = 0$ . This quantity is shown as the second y-axis in all relevant figures in this work.

The various baryonic phases are distinguished in different ways in the simulations. For L-GALAXIES 2020, baryons within haloes are split into seven distinct components: SMBH, bulge stars, disc stars, halo stars, cold gas (i.e. ISM), hot gas (i.e. CGM), and an ejecta reservoir. The cold gas component within galaxies (which is tuned to reproduce the observed  $\text{H}$  mass function at  $z = 0$ , see [Henriques et al. 2020](#)) is assumed to represent neutral gas in this work. For EAGLE and TNG100, gas particles/cells are assigned to phases according to their volumetric density. Density ranges have



**Figure 1.** Evolution of the cosmic metal density in various phases in and around galaxies from  $z = 7$  to 0. Each panel represents one of the four galaxy evolution models considered here (see legends). Lines represent model results, with green shaded regions representing the uncertainty in the DLA density threshold for EAGLE and TNG100 (see Section 4.1.1). Points (shown in all four panels) represent observational results from PH20. Dotted green lines represent  $\rho_{\text{met,neu}}$  for L-GALAXIES 2020 when corrected with the HI density model described in Section 4.3. Dashed green line represents  $\rho_{\text{met,neu}}$  when only considering galaxies with  $\log_{10}(M_*/M_{\odot}) > 7.5$  in L-GALAXIES 2020 MM (see Section 4.4).

been chosen to roughly replicate the column density designations given to DLAs, sub-DLAs, LLSs, and sub-LLSs in observations (see Fig. 1). Therefore, dense neutral gas is assumed to be represented by material with  $n_{\text{gas}} > 10^{-1.3} \text{ cm}^{-3}$ . This corresponds to the threshold column density for DLAs of  $N(\text{HI}) = 10^{20.3} \text{ cm}^{-2}$  at  $z = 3$  in the EAGLE precursor simulations studied by Rahmati et al. (2013) (see also PH20, fig. 1c). Because the relation between volumetric and column densities is not exact, a factor of 2 uncertainty (i.e. a range of thresholds from  $10^{-1.6}$  to  $10^{-1.0} \text{ cm}^{-3}$ ) is accounted for in this work, matching the  $1\sigma$  spread in  $n_{\text{gas}}$  found by Rahmati et al. (2013) at  $N(\text{HI}) = 10^{20.3} \text{ cm}^{-2}$ . This uncertainty in EAGLE and TNG100 results is indicated by shaded regions in all relevant figures.

#### 4.1.2 Apportionment of metals

Fig. 1 shows the cosmic metal density evolution for various phases in and around galaxies. Each panel shows results from a different simulation considered here (lines), with PH20 observational data shown in all panels (points).

At low redshift ( $z \lesssim 2.5$ ), all four simulations reproduce the observed partition of metals among various phases reasonably well. For example, EAGLE and L-GALAXIES 2020 (both the DM and MM) match the steep increase in  $\rho_{\text{met,*}}$  observed from  $z \sim 0.75$  to the present day (orange). EAGLE and TNG100 also roughly reproduce the normalisation of the metal density in neutral gas at  $z \sim 0.5$  from DLAs (green), as well as the total gas metal density at  $z \sim 1$  inferred from SED modelling of present-day galaxies (Bellstedt et al. 2020).

Importantly, the metals found in gas with  $n_{\text{gas}} \leq 10^{-2.4} \text{ cm}^{-3}$  in EAGLE and TNG100, or equivalently in the hot CGM in L-

GALAXIES 2020, make up a significant fraction of the total metal budget (solid red lines). The dominance of this phase in the simulations highlights its importance when assessing the “missing metals problem”. Such material is roughly equivalent to sub-LLS gas with  $N(\text{H}) < 1.6 \times 10^{17} \text{ cm}^{-2}$  in observations, as traced by the Ly $\alpha$  forest. The full complement of this low-density phase is difficult to measure observationally, especially at high redshift, so is not included in the observational data set shown in Fig. 1. However, the sub-component of this phase from the IGrM+ICM at low-redshift is shown (red points). This can be compared to the dashed red lines from the simulations, which represents hot gas in within subhaloes of mass  $\log(M_{200}/M_{\odot}) \geq 13.0$ . In all three simulations, this sub-component has a relatively small contribution to the total low-density gas in the universe, which is predominantly found in the subhaloes of field galaxies. In TNG100, the dominant metal fraction is shared roughly equally between this low-density gas and stars, highlighting the large amount of metals locked-up in stars in that simulation.

At high redshift ( $z \gtrsim 2.5$ ), the differences between the simulations and observations are more pronounced. For example, EAGLE appears to under-estimate  $\rho_{\text{met,neu}}$  by 0.27 dex at  $z \sim 4$  compared to observations, whereas the L-GALAXIES 2020 DM and MM under-estimate this much more significantly, by 0.49 dex and 0.82 dex, respectively. In turn, TNG100 is able to reasonably match the observed  $\rho_{\text{met,neu}}$  from DLAs at high redshift. However, in order to do this, the total metal budget (black line), is higher than that inferred from either observations or the other simulations considered here. The cause of this discrepancy is discussed in Section 4.2.

Overall, the observational claim that nearly all metals at high redshift are in neutral gas is not reproduced in any of the simulations. While DLA observations suggest that  $\sim 87$  per cent of the total expected metal budget at  $z \sim 4$  is in neutral gas, less than 40 per cent is in neutral gas in the simulations considered here at the same redshift (19 per cent for L-GALAXIES 2020 DM, 8 per cent for L-GALAXIES 2020 MM, 30 per cent for EAGLE, and 32 per cent for TNG100). Hot gas, and stars in the case of TNG100, also have a major contribution in the simulations. This material remains largely unobserved at these redshifts.

#### 4.1.3 Biases in observations and simulations

When considering the accuracy of  $\rho_{\text{met,neu}}$  in observations, we note that the dust corrections applied to  $\langle M/H \rangle_{\text{DLA}}$  measurements are significant (up to  $\sim 0.5$  dex for iron at high redshift), and can, in principle, lead to over-estimates. However, the dust correction technique utilised by De Cia et al. (2018) for this data set is relatively well constrained, relying on multiple metal absorption lines. This means that the dust corrections applied here are likely to be reliable. A lack of highly dust-obscured systems in the sample can also pose an issue, however this would lead to an *under*-estimate of  $\rho_{\text{met,neu}}$ , rather than an over-estimate, therefore *increasing* the discrepancy between the observations and simulations. Dust is therefore unlikely to fully explain the discrepancy seen in Fig. 1.

When considering the accuracy of metal density measurements in simulations, we note that mass resolution limits could lead to inaccurate estimates for very low mass systems (see e.g. Di Gioia et al. 2020). However, the amount of metals residing within sub-resolution systems would have to be considerably under-estimated for this to fully account for the discrepancies seen in Fig. 1. Nonetheless, such mass resolution issues do significantly affect model H I mass densities, as discussed in Section 4.3.

Finally, in contrast to the argument that model galaxies *below* a certain mass should contribute more metals, an alternative argument

is that the observed DLA samples are only probing systems *above* a certain mass. We find that when limiting the minimum stellar mass assumed for DLA hosts in the simulations (and therefore using Eqn. 2 rather than Eqn. 1 to calculate  $\rho_{\text{met,neu}}$ ), the correspondence with the observed  $\rho_{\text{met,neu}}$  from DLAs is much improved for L-GALAXIES 2020 MM (green dashed line in Fig. 1). We discuss this possibility and other complications in the estimation of  $\rho_{\text{met,neu}}$  from DLAs in Section 4.4.

In conclusion, we find that none of the cosmological simulations considered here agree with the observational finding that the majority of metals in the Universe at high redshift are in the neutral gas phase. Instead, all four simulations suggest that low-density hot gas also has a significant contribution. However, the absolute amount of metals observed in DLAs can be matched by simulations, but only under certain conditions outlined in the following sections.

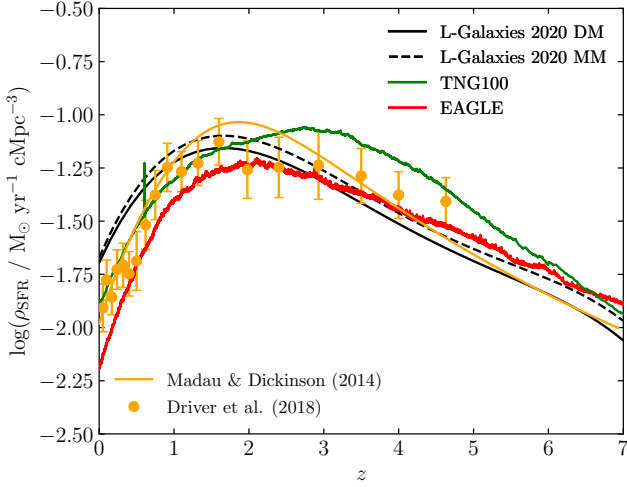
## 4.2 Star formation rate density evolution

When assessing the metal densities seen at high redshift, it is important to also consider the evolution of the cosmic SFR density ( $\rho_{\text{SFR}}$ ). This is because, to first order,  $\rho_{\text{SFR}}$  controls the rate of metal production in the Universe and the distribution of these metals throughout galaxies and their environments via outflows.

Fig. 2 shows the evolution of  $\rho_{\text{SFR}}$  for the four simulations considered here, along with two observational compilations from Madau & Dickinson (2014) (orange line) and Driver et al. (2018) (orange points). Both these studies incorporate data spanning rest-frame wavelengths from the far infrared (FIR, at lower redshifts) to the far ultraviolet (FUV). Comparing the observed  $\rho_{\text{SFR}}$  with theoretical models in this way is difficult, as many systematic uncertainties exist, including the importance of obscured star formation, biased galaxy samples, resolution effects, and cosmic variance. Nonetheless, we present here a face-value comparison, and discuss the complexities involved.

The L-GALAXIES 2020 DM and MM match the observed  $\rho_{\text{SFR}}$  shown in Fig. 2 quite well at high redshift, but over-estimate the total amount of star formation below  $z \sim 1$  (when run on the MILLENNIUM-II). Conversely, EAGLE slightly over-estimates the observed  $\rho_{\text{SFR}}$  at  $z \gtrsim 5$ , but under-estimates it by over 0.2 dex at  $z \lesssim 2$  (see also Furlong et al. 2015; Katsianis et al. 2017; Driver et al. 2018). TNG100 has comparable (or perhaps too *low*, Donnari et al. 2019) star formation rates at  $z \sim 2$ , and is also consistent with 3D-HST data at  $z \sim 1.5$  (Nelson et al. 2021). However, TNG100 also shows the largest difference at high redshift, with significantly more star formation at  $3 \lesssim z \lesssim 5$  than inferred by Madau & Dickinson (2014) and Driver et al. (2018), and an earlier peak at  $z \sim 3$  (see also Bellstedt et al. 2020; Zhao et al. 2020). Cosmic stellar mass densities,  $\rho_{*}$ , are also higher than observed. This enhanced star formation relative to the other simulations at  $z \sim 3 - 5$  is the cause of the higher metal densities seen in TNG100 in Fig. 1.

Furlong et al. (2015) discuss how a combination of burstier star formation and reduced AGN feedback efficiency would enable higher SFRs and therefore increased  $\rho_{\text{SFR}}$  in EAGLE, particularly around the peak of cosmic star formation at  $z \sim 2$ . However, we note that this could also lead to an over-estimate of  $\rho_{\text{met,neu}}$  compared to low-redshift DLA observations, which are already slightly exceeded by EAGLE (see Fig. 1). Similarly, a reduction in the star formation efficiency at high redshift in TNG100 (particularly in the progenitors of the most-massive galaxies by  $z = 0$ , see Bellstedt et al. 2020) would lead to a better agreement with current  $\rho_{\text{SFR}}$  inferences from FUV data, but would also cause a greater under-estimate of  $\rho_{\text{met,neu}}$  at high redshift.



**Figure 2.** Cosmic star formation rate density versus redshift for the four galaxy evolution models considered here. Observational data from [Madau & Dickinson \(2014, orange line\)](#) and [Driver et al. \(2018, orange points\)](#) are also shown for comparison. We contrast four simulations: L-GALAXIES 2020 SM (solid black), L-GALAXIES 2020 MM (dashed black), TNG100 (green), and EAGLE (red).

Therefore, a comparison between Fig. 1 and Fig. 2 suggests an incompatibility between the currently observed  $\rho_{\text{met,neu}}$  and  $\rho_{\text{SFR}}$  from FIR-FUV data. This incompatibility is reminiscent of that noted between some  $\rho_{\text{SFR}}$  and  $\rho_*$  estimates, which is now largely understood (e.g. [Wilkins et al. 2019](#); [Leja et al. 2020](#); [Koushan et al. 2021](#)).

When assessing this incompatibility, we note that current observational determinations of  $\rho_{\text{SFR}}$  using FUV data could underestimate the amount of star formation in the high-redshift Universe (e.g. [Bellstedt et al. 2020](#); [Asada et al. 2021](#)). Significant uncertainties in  $\rho_{\text{SFR}}$  are caused by limited sky coverage and the incompleteness of samples, which make large, uncertain extrapolations of the galaxy luminosity function necessary (e.g. [Madau & Dickinson 2014](#); [Katsianis et al. 2017](#)). Dusty star-forming galaxies can also be missed when selecting and/or analysing galaxies at shorter wavelengths (e.g. [Rowan-Robinson et al. 2016](#); [Wang et al. 2019](#)). Recent studies of ALMA data for a small number of systems from the MORA and ALPINE surveys suggest that such obscured star formation could contribute between 17 and 35 per cent of the total  $\rho_{\text{SFR}}$  at  $z \sim 5$  (e.g. [Gruppioni et al. 2020](#); [Zavala et al. 2021](#)). Inferences from GRBs also suggest a higher  $\rho_{\text{SFR}}$  at  $z > 5$  ([Kistler et al. 2009](#)), and a recent study by [Matthews et al. \(2021\)](#) utilising radio data suggests a systematically higher  $\rho_{\text{SFR}}$  at  $z \lesssim 3.5$ . Importantly, at  $z \sim 4$ , where TNG100 shows the biggest discrepancy with observations in Fig. 2, [Katsianis et al. \(2020\)](#) find that rest-frame UV data could under-predict SFRs in more massive galaxies by  $\sim 0.13$  dex. Although, we note that other studies have shown that FIR-FUV measurements can instead *over-estimate* the true  $\rho_{\text{SFR}}$  (e.g. [Utomo et al. 2014](#); [Katsianis et al. 2017](#); [Leja et al. 2019](#); [Katsianis et al. 2020](#)).

A systematic bias in the observed  $\rho_{\text{SFR}}$  will also affect the total expected metal density,  $\rho_{\text{met,tot}}$  (black points in Fig. 1). [PH20](#) determine this quantity by integrating the  $\rho_{\text{SFR}}$  from [Madau & Dickinson \(2014\)](#) over time, assuming a fixed mass return fraction and metal yield. It is therefore sensitive not only to changes in stellar lifetimes, remnant masses, and metal yields with redshift, but also the accuracy of SFR measurements. Therefore, an increase in  $\rho_{\text{SFR}}$  at  $3 \lesssim z \lesssim 5$  would cause a corresponding increase in  $\rho_{\text{met,tot}}$ , reducing the fraction of

total metals found in neutral gas. This would reduce the discrepancy between observations and simulations discussed in Section 4.1.

However, in addition to observational uncertainties, biases in simulations due to their limited resolution and volumes should also be considered ([Schaye et al. 2010](#); [Pillepich et al. 2018b](#)). We find that the  $\rho_{\text{SFR}}$  obtained from the larger TNG300 simulation is  $\sim 0.2$  dex *lower* than seen in TNG100, bringing it into closer agreement with FIR-FUV observations at  $z \sim 3$ , although worse agreement below  $z \sim 2$  (see also [Zhao et al. 2020](#)). Similarly,  $\rho_{\text{SFR}}$  in L-GALAXIES 2020 when run on the larger MILLENNIUM-I simulation is lower than when run on MILLENNIUM-II (see also [Henriques et al. 2020](#)). These differences are predominantly due to lower resolution in the larger simulations, rather than cosmic variance issues in the smaller simulations ([Genel et al. 2014](#); [Zhao et al. 2020](#)). Lower resolution leads to an under-abundance of low-mass galaxies, which in turn leads to an under-estimation of  $\rho_{\text{HI}}$  and  $\rho_{\text{SFR}}$ . For example, [Spinelli et al. \(2020\)](#) find that haloes with  $M_{200} \lesssim 10^{11} M_{\odot}$  contribute the majority of  $\rho_{\text{HI}}$  above  $z \sim 2.5$  in GAEA (see also [Lagos et al. 2014](#)), and [Villaescusa-Navarro et al. \(2018\)](#) find an under-estimate in  $\rho_{\text{HI}}$  at high redshift in TNG300 relative to TNG100. Likewise,  $\rho_{\text{SFR}}$  in both ILLUSTRIS and EAGLE is dominated by lower-mass galaxies at high redshift ([Genel et al. 2014](#); [Katsianis et al. 2017](#)).

We therefore conclude that one avenue for cosmological simulations to reproduce the observed  $\rho_{\text{met,neu}}$  at high redshift is to produce stars (and therefore metals) at a higher rate than is expected from current FIR-FUV observations of  $\rho_{\text{SFR}}$ . This implies a tension between the currently-observed  $\rho_{\text{met,neu}}$  and  $\rho_{\text{SFR}}$ , and highlights the importance of considering both properties when studying galaxy evolution.

### 4.3 Neutral gas density evolution

While the cosmic  $\rho_{\text{SFR}}$  plays an important role in the intrinsic evolution of the metal density in neutral gas, measurements of  $\rho_{\text{HI}}$  are equally important in its observational estimation via Eqn. 2. At  $z \leq 0.4$ , cosmic  $\rho_{\text{HI}}$  is typically inferred via measurements of  $M_{\text{HI}}$  from the 21-cm emission line. An HI mass function is then formed and integrated between reasonable limits to obtain  $\rho_{\text{HI}}$ . At  $0.4 \lesssim z \lesssim 5.0$ , measurements of the Lyman- $\alpha$  absorption line in the spectra of background quasars are used to determine the HI column density,  $N_{\text{HI}}$ . The column density distribution function (CDDF) can then be formed, from which  $\rho_{\text{HI}}$  is derived in a similar way to that at lower redshifts. [PH20](#) provide a linear fit to the redshift evolution of  $\rho_{\text{HI}}$  obtained from a compilation of recent surveys utilising these two techniques, given by,

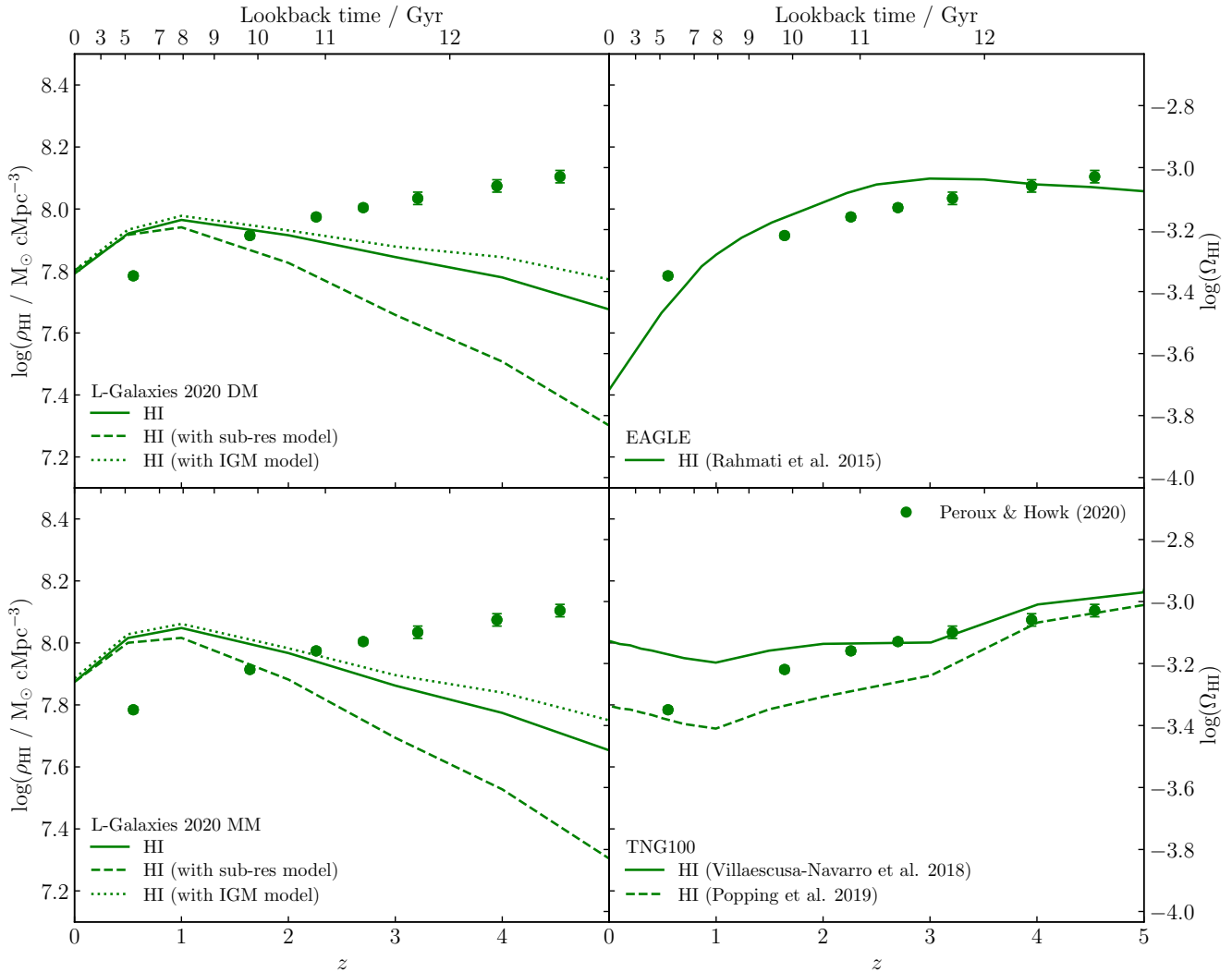
$$\rho_{\text{HI,obs}}(z) = \rho_{\text{crit,0}} (4.6_{\pm 0.2} \times 10^{-4}) (1+z)^{0.57_{\pm 0.04}} - (1/0.76) \quad , \quad (3)$$

where the correction factor of  $1/0.76$  accounts for helium. The green points in Fig. 3 represent this fit, plotted at the discrete redshifts of the binned DLA data used in Sections 4.1 and 4.4. The recent analysis of [Heintz et al. \(2021\)](#), which utilises [CII] emission line measurements and a [CII] – HI conversion factor calibrated to GRB host galaxies, found a similar rate of increase in  $\rho_{\text{HI}}$  with redshift above  $z \sim 2$ , with this trend flattening somewhat at  $z \sim 5 - 6$ .

For the L-GALAXIES 2020 models, the HI density shown in Fig. 3 is calculated from the total hydrogen mass in cold gas,  $M_{\text{H}}$ , as,

$$\rho_{\text{HI,theo}} = \frac{\sum (1 - f_{\text{H2}}) M_{\text{H}}}{V_{\text{tot}}} \quad , \quad (4)$$

where  $V_{\text{tot}}$  is again the volume of the simulation box, and  $f_{\text{H2}}$



**Figure 3.** Evolution in HI mass density in galaxies from  $z = 5$  to 0. Each panel corresponds to a different galaxy evolution simulation. For the L-GALAXIES 2020 simulations (left panels), solid lines represent default results, dashed lines represent  $\rho_{\text{HI}}$  when assuming a separate model for the HI mass in sub-resolution systems, and dotted lines represent  $\rho_{\text{HI}}$  when assuming a separate model for HI mass in the IGM (see Section 4.3). For TNG100 (bottom right), the solid and dashed lines show complementary measurements of  $\rho_{\text{HI}}$  from two different works (see legend). In all panels, points denote observational data from PH20.

is the molecular hydrogen fraction as calculated following the McKee & Krumholz (2010) gas partition formalism (see Henriques et al. 2020, section 2.2.3). For EAGLE,  $\rho_{\text{HI}}$  is taken directly from a previous analysis by Rahmati et al. (2015). In that study, the equilibrium HI fraction in each gas particle is calculated by balancing the collisional and UVB photoionization rates (as parameterised by Haardt & Madau 2001) with the Case A recombination rate, assuming a temperature of  $10^4$  K for star-forming gas and self shielding in denser regions (see Rahmati et al. 2013). For TNG100, cosmic  $\rho_{\text{HI}}$  is taken from two complementary analyses by Villaescusa-Navarro et al. (2018) and Popping et al. (2019), which applied two different prescriptions for HI – H<sub>2</sub> gas partitioning. Villaescusa-Navarro et al. (2018) follow a similar prescription to Rahmati et al. (2015), but adopt the UVB model of Faucher-Giguère et al. (2009). Popping et al. (2019) instead first calculate the H<sub>2</sub> mass following the metallicity-dependent formalism of Gnedin & Kravtsov (2011), then remove this from the total neutral gas budget to obtain the HI mass.

We highlight here that these recipes also account for the presence of HI in low-density, non-star-forming gas in EAGLE and TNG100.

Such gas is not modelled in L-GALAXIES 2020, contributing to the significantly lower  $\rho_{\text{HI}}$  predicted at high redshift in that simulation, as discussed below.

The trend of increasing  $\rho_{\text{HI}}$  with increasing redshift seen in observations is well reproduced by EAGLE and TNG100, as is shown in previous works (e.g. Rahmati et al. 2015; Villaescusa-Navarro et al. 2018; Diemer et al. 2019). However, we note that the observed densities inferred from Lyman- $\alpha$  absorption at higher redshift directly probe only dense DLA gas with  $N_{\text{HI}} \geq 10^{20.3} \text{ cm}^{-2}$ , with an extrapolation down to lower column densities required (typically down to  $N_{\text{HI}} \sim 10^{12} \text{ cm}^{-2}$ , see Zafar et al. 2013; Berg et al. 2019). Less than 40 per cent of the total HI mass is actually at DLA densities in TNG100, at all redshifts, and the evolution of this dense neutral gas is flatter than shown for all HI gas in Fig. 3 (see Hassan et al. 2020).

L-GALAXIES 2020 only reproduces the observed trend in  $\rho_{\text{HI}}$  below  $z \sim 1$ , albeit with a higher normalisation than observations (while still matching the observed HI mass function from Zwaan et al. 2005 at  $z = 0$ , see Yates et al. 2021). Above  $z \sim 2$ , both L-GALAXIES 2020 models suggest a *decrease* in  $\rho_{\text{HI}}$  with redshift, in contrast to



observations. This discrepancy has also been seen in earlier versions of L-GALAXIES (Martindale et al. 2017), as well as the GALFORM, SANTA CRUZ, and GAEA semi-analytic models (Lagos et al. 2011; Berry et al. 2014; Spinelli et al. 2020), and the fiducial analytic model presented by Theuns (2021). These works suggest that this  $\rho_{\text{HI}}$  deficit is due to a lack of adequate accounting for HI in sub-resolution haloes and/or neutral gas outside galaxies. The modelling of the ultraviolet background (UVB) radiation and of  $\text{H}_2$  could also play a role, although we note that the uniform UVB models adopted by the simulations considered here predict a similar evolution in the HI photoionisation rate with redshift (see Faucher-Giguère 2020).

When considering HI in sub-resolution haloes, we note that Di Gioia et al. (2020) utilised a halo occupation distribution (HOD) in GAEA to more accurately model the number and HI content of such systems. We mimic this approach here for L-GALAXIES 2020, by taking model HI measurements only from resolved subhaloes with masses above  $M_{200} = 10^{9.2} M_{\odot}$  and adding a fixed contribution from lower-mass systems of  $1.5 \times 10^{-5} \rho_{\text{crit}} M_{\odot}/\text{Mpc}^3$  at all redshifts (see Di Gioia et al. 2020, fig. 6b). This evolution is shown by the dashed green lines in the left-hand panels of Fig. 3. As found by Spinelli et al. (2020) and Di Gioia et al. (2020), this approach still under-estimates the HI densities observed at  $z \gtrsim 3$ . Indeed, the HOD-based prescription actually predicts lower cosmic HI densities at high redshift than are obtained when including the HI masses from sub-resolution haloes in L-GALAXIES 2020 (solid green lines). Di Gioia et al. (2020) found that an *a posteriori* increase in the HI content of *all* galaxies by a factor of two was also required before  $\rho_{\text{HI}}$  was significantly increased in GAEA. This suggests that corrections to the gas content of sub-resolution systems alone is not enough to resolve the discrepancy in HI density seen between semi-analytic models and observations.

When considering HI outside of galaxies, we note that Villaescusa-Navarro et al. (2018) find  $\sim 6$  per cent of the cosmic HI at  $z = 3$  in TNG100 lies beyond the subhalo boundary, with this percentage increasing to 20 per cent by  $z = 5$ . Such a neutral intergalactic medium (IGM) is not modelled at all in semi-analytic models (see also Lagos et al. 2014; Xie et al. 2017). Therefore, we mimic the effect of including this IGM HI gas, by adding its fraction as seen in TNG100 (see Villaescusa-Navarro et al. 2018, fig. 3) to that found inside galaxies in L-GALAXIES 2020. This adapted  $\rho_{\text{HI}}$  evolution is shown by the dotted green lines in the left-hand panels of Fig. 3. While this correction does increase  $\rho_{\text{HI}}$  in L-GALAXIES 2020, especially at higher redshift, it still falls short of that observed at  $z \gtrsim 3$ . This indicates that larger HI masses *within resolved subhaloes* are still required in semi-analytic models at high redshift.

This missing neutral gas should presumably lie in the outskirts of galaxies and/or in the inner CGM. We find that a substantial amount ( $\sim 65$  per cent) of the HI gas within subhaloes in TNG100 lies beyond  $2r_e$  at  $z \gtrsim 3$ , where  $r_e$  is the stellar half-mass radius (see Byrohl et al. 2021, for observable implications). A similar conclusion has been drawn for the EAGLE simulation by Garratt-Smithson et al. (2021), who find up to 50 per cent of HI gas within massive subhaloes is in the CGM at  $z = 2$ . Likewise, Stevens et al. (2019) point to the importance of the CGM in TNG100 when comparing HI masses to wide-aperture observations of the 21cm line at lower redshift.

Given this, we make a simple test to see if an *ad hoc* correction to the  $\rho_{\text{HI}}$  within subhaloes in L-GALAXIES 2020 could also resolve the discrepancy in  $\rho_{\text{met,neu}}$  at high  $z$  in Fig. 1. We do this by assuming that the same deficit in HI seen between L-GALAXIES 2020 and observations in Fig. 3 is also present for metals, such that the mean cosmic cold gas metallicity in the simulation is unchanged. The

adapted  $\rho_{\text{met,neu}}$  evolution that this simple test produces is shown as green dotted lines in the left-hand panels of Fig. 1.

At  $z \lesssim 2$ , the lower  $\rho_{\text{met,neu}}$  caused by this correction (due to the apparent *excess* of HI in the models) brings both the DM and MM into improved agreement with DLA observations. At  $z \gtrsim 2$ , this *ad hoc* correction also nicely improves agreement for L-GALAXIES 2020 DM. However, this version of the simulation is already known to over-estimate ISM metallicities in high- $z$  galaxies (see Yates et al. 2021). For L-GALAXIES 2020 MM, which reproduces ISM metallicities in high- $z$  galaxies well (see Section 4.4),  $\rho_{\text{met}}$  remains too low compared to DLA observations when assuming an unchanged mean cosmic cold gas metallicity. Instead, we find that a mean metallicity for the added gas of  $12 + \log(\text{O}/\text{H}) \sim 8.45$  is required at  $z \sim 4$  to fully resolve the  $\rho_{\text{met,neu}}$  discrepancy seen, which is roughly the metallicity of galaxies with  $\log_{10}(M_*/M_{\odot}) = 10.5$  at that redshift in L-GALAXIES 2020 MM.

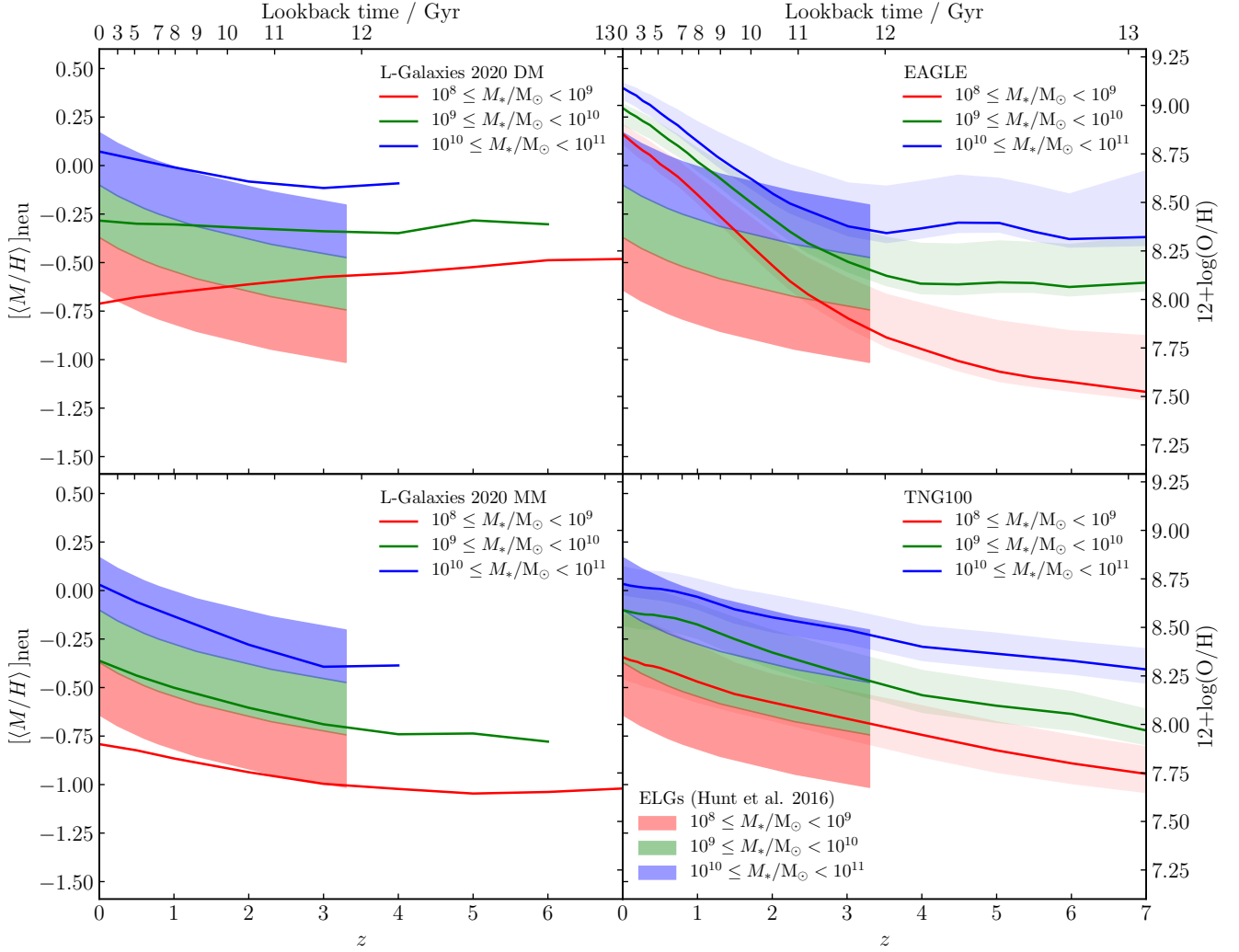
Such high-metallicity material would require a significant increase in  $\rho_{\text{SFR}}$ , and consequently  $\rho_*$ , at high redshift, which may bring L-GALAXIES 2020 into better agreement with TNG100 (although worse agreement with observations from Madau & Dickinson 2014). However, this outcome is far from certain, as an increase in cold gas mass would also necessitate changes to the rates of gas cooling,  $\text{H}_2$  formation, and other processes in L-GALAXIES 2020. Clearly, further study of the implications (and physical motivations) for such a change in semi-analytic models is required before a more definitive statement can be made. Such an HI-based correction would also do little to resolve the smaller deficit in  $\rho_{\text{met}}$  seen in EAGLE, which reproduces  $\rho_{\text{HI}}$  reasonably well at high redshift.

We therefore conclude that a simple increase in the amount of neutral gas in or around galaxies in simulations is unlikely to provide a complete solution to the neutral gas metal discrepancy seen in Fig. 1. However, it is likely to be a contributing factor in the case of semi-analytic models, in combination with the other factors discussed in Section 4.4.

#### 4.4 Neutral gas metallicity evolution

In addition to the neutral gas density discussed in Section 4.3, the mean neutral gas metallicity,  $[\langle \text{M}/\text{H} \rangle]_{\text{neu}}$ , also plays an important role in determining  $\rho_{\text{met,neu}}$  in observations via Eqn. 2. Fig. 1 shows that the evolution in  $\rho_{\text{met,neu}}$  as probed by DLAs is flatter than that seen in galaxy evolution simulations (and in observational  $\rho_{\text{met,tot}}$  estimates, see also Rafelski et al. 2014). However, the corresponding evolution in  $[\langle \text{M}/\text{H} \rangle]_{\text{neu}}$  for DLAs is actually *steeper* than seen in simulations and other ISM observations *at fixed mass*. For example, De Cia et al. (2018) find an evolution in  $[\langle \text{M}/\text{H} \rangle]_{\text{neu}}$  of  $\sim 0.63$  dex from  $z \sim 3.2$  to 0.5 for their DLA sample. In contrast, estimates of the mean ISM oxygen abundance evolution at fixed mass from recent emission-line studies (using direct and indirect metallicity diagnostics) do not exceed  $\sim 0.38$  dex across the same redshift interval (e.g. Izotov et al. 2015; Hunt et al. 2016; Sanders et al. 2016, 2021). Studies which also utilise GRB absorption-line metallicities at high redshift (which, like DLAs, should be unbiased in terms of host galaxy luminosity) also find similar results (e.g. Thöne et al. 2013; Graham et al. 2019; Yates et al. 2021). In several recent cosmological simulations, the evolution in the ISM oxygen abundance of star-forming galaxies at fixed mass is also found to be only 0.25–0.45 dex from  $z \sim 3$  to 0 (see Yates et al. 2021 section 4.2, and references therein).

One possible explanation for this discrepancy is that the average mass of DLA host galaxies evolves with time. This may seem unlikely, as one of the main strengths of DLAs is that their detectability



**Figure 4.** Evolution of gas-phase metallicity in three fixed-mass bins (see legend). Observational results for ELGs from  $z = 0$  to 3.3 by [Hunt et al. \(2016\)](#) are shown as dark shaded regions. The mean, HI-weighted gas metallicities for model galaxies in the four simulations considered here are also shown. Light shaded regions represent the uncertainty in the DLA density threshold for EAGLE and TNG100 (see Section 4.1.1). All metallicities are normalised to the solar photospheric value of  $M_{\text{met},\odot}/M_{\text{H},\odot} = 0.0181$  from [Asplund et al. \(2009\)](#).

is not biased towards brighter (i.e. more massive) systems at any particular redshift. Typical optical and UV sensitivities are also well below that required to detect the metal absorption lines seen in current DLA samples (see [PH20](#), section 3.4). Nonetheless, the masses of DLA host galaxies are currently poorly constrained. For example, only  $\sim 5$  of the 235 DLAs in the [De Cia et al. \(2018\)](#) sample have identified galaxy counterparts with measured stellar masses and impact parameters ([Krogager et al. 2013](#); [Fynbo et al. 2013](#); [Christensen et al. 2014](#); [Augustin et al. 2018](#); [Rhodin et al. 2021](#)).

Therefore, in this section we attempt to quantify a possible mass dependence of  $[\langle M/H \rangle]_{\text{neu}}$  measurements from a theoretical perspective. To do this, metallicities from the simulations are measured in an equivalent way to the DLA observations, by taking the log of the HI-weighted mean metal-to-hydrogen mass ratio in cold, dense gas,

$$[\langle M/H \rangle]_{\text{neu}} = \log_{10} \left[ \frac{\sum (M_{\text{met}}/M_{\text{H}}) M_{\text{HI}}}{\sum M_{\text{HI}}} \right] - \log_{10} \left[ \frac{M_{\text{met},\odot}}{M_{\text{H},\odot}} \right], \quad (5)$$

with both model and observational values normalised to the solar photospheric value of  $Z_{\odot}/X_{\odot} = M_{\text{met},\odot}/M_{\text{H},\odot} = 0.0181$  from [As-](#)

[plund et al. \(2009\)](#). The total metal and hydrogen masses in the cold, dense gas of each model galaxy are used, such that  $[\langle M/H \rangle]_{\text{neu}}$  represents the cosmic average metallicity for galaxies. Weighting each galaxy’s contribution by their HI mass mitigates somewhat any differences in the fraction of star-forming galaxies among the various observational and model samples.

#### 4.4.1 Metallicity evolution at fixed mass

Before comparing directly with DLA metallicities, we first show a comparison between simulations and the expected ISM metallicity evolution in star-forming emission-line galaxies (ELGs) as probed by composite spectra dominated by HII regions. Fig. 4 shows the evolution in  $[\langle M/H \rangle]_{\text{neu}}$  from L-GALAXIES 2020, EAGLE, and TNG100, as defined by Eqn. 5, for three, 1-dex-wide, fixed-mass bins (solid lines), alongside the evolution from a compilation of ELG samples by [Hunt et al. \(2016\)](#) (dark shaded regions). These observations were re-calibrated to the [Pettini & Pagel \(2004\)](#) N2 strong-line diagnostic, which returns good agreement with various direct metallicity mea-

measurements at both low and high redshift (Andrews & Martini 2013; Yates et al. 2021). However, we acknowledge that other strong-line diagnostics predict higher metallicities at low redshift (e.g. Maiolino et al. 2008; Zahid et al. 2014).

The evolution in model H $\alpha$ -weighted neutral gas metallicities shown here is similar to that previously reported when using SFR-weighted methods and oxygen abundances in each of the simulations considered (see De Rossi et al. 2017 for EAGLE, Torrey et al. 2019 for TNG100, and Yates et al. 2021 for L-GALAXIES 2020). Although, we note that plotting the average  $[\langle M/H \rangle]_{\text{neu}}$  in 1-dex-wide mass bins here naturally leads to a slightly steeper evolution than would be seen for an exact fixed mass, due to the increasing contribution to cosmic metals (as well as cosmic H $\alpha$  and SFR, see Lagos et al. 2014; Genel et al. 2014; Spinelli et al. 2020) from higher-mass, more metal-rich galaxies over cosmic time, at least for systems below  $\log_{10}(M_*/M_{\odot}) \sim 10.5$ .

Fig. 4 shows that only L-GALAXIES 2020 MM is able to reproduce the fixed-mass metallicity evolution for ELGs observed by Hunt et al. (2016) back to  $z \sim 3.3$  (with the exception of the lowest-mass galaxies at the lowest redshifts). L-GALAXIES 2020 DM returns no clear evolution in  $[\langle M/H \rangle]_{\text{neu}}$  at all, due to an over-retention of newly-synthesised metals in the ISM at early times (see Yates et al. 2021). TNG100 returns a higher normalisation than observed by Hunt et al. (2016), and EAGLE has both a higher normalisation and steeper slope. Although, we note that the Recal-L025N0752 version of EAGLE returns lower metallicities in low-mass galaxies than the Ref-L100N1504 version used here, due to its stronger SN feedback (Schaye et al. 2015; De Rossi et al. 2017). We also find that, at high redshift, gas with  $n_{\text{gas}} > 10^{-1.0} \text{ cm}^{-3}$  in EAGLE (i.e. the upper limit of the uncertainty bands) appears to have much higher metallicity, by  $\gtrsim 0.2$  dex, than gas with  $n_{\text{gas}} > 10^{-1.3} \text{ cm}^{-3}$  (solid lines), indicating a strong dependence of metallicity on gas density. This has also been identified in EAGLE by De Rossi et al. (2017) and Garratt-Smithson et al. (2021).

In general, low-mass galaxies are much more metal rich by  $z = 0$  in EAGLE and TNG100 than in L-GALAXIES 2020. The Ref-L100N1504 EAGLE simulation even predicts super-solar ISM metallicities in galaxies with masses as low as  $\log_{10}(M_*/M_{\odot}) \sim 8$  at  $z = 0$ . This is predominantly due to differences in the SN feedback efficiency in these simulations. As discussed in Section 3.4, the reference EAGLE simulation shown here has relatively low SN feedback efficiency, leading to higher metallicities in lower-mass systems (Schaye et al. 2015; De Rossi et al. 2017). This provides reasonable agreement with older, indirect ISM metallicity measurements for ELGs at low redshift (e.g. Tremonti et al. 2004; Kobulnicky & Kewley 2004). On the other hand, the increased metal ejection efficiency from SNe in L-GALAXIES 2020 MM leads to better agreement with more recent indirect and direct metallicity measurements (e.g. Hunt et al. 2016; Dopita et al. 2016; Yates et al. 2020).

Information on the gas-phase mass – metallicity relation (MZ $_{\text{gR}}$ ) at various redshifts is also obtainable from Fig. 4. For example, at  $z \sim 3 - 4$ , we can see that the MZ $_{\text{gR}}$  in L-GALAXIES MM is steeper than that in EAGLE (i.e. the separation between the fixed-mass lines in Fig. 4 is greater), which in turn is steeper than that in TNG100. This ordering is reflective of the relative  $\rho_{\text{met,neu}}$  in these simulations at these redshifts, as shown in Fig. 1, demonstrating the importance of the metal content in lower-mass galaxies. We further note that, for gas with  $n_{\text{gas}} > 10^{-1.0} \text{ cm}^{-3}$ , the EAGLE MZ $_{\text{gR}}$  becomes as flat as that in TNG100 at  $z \sim 3 - 4$ , due to the significant presence of metals in the densest regions.

#### 4.4.2 Mass dependencies in DLA host samples

Fig. 5 shows the observed evolution in H $\alpha$ -weighted  $[\langle M/H \rangle]_{\text{neu}}$  for the De Cia et al. (2018) DLA sample (black circles) from  $z \sim 5$  to 0. Solid lines denote the corresponding evolution for galaxies in the four simulations considered here, for three fixed lower-mass thresholds (see legend). A clear evolution in  $[\langle M/H \rangle]_{\text{neu}}$  is seen, in contrast to earlier determinations using smaller samples and relatively under-depleted metallicity indicators such as zinc (e.g. Vladilo et al. 2000).

When considering galaxies without any lower-mass threshold (red lines), the  $[\langle M/H \rangle]_{\text{neu}}$  evolution seen in simulations is indeed steeper than seen at fixed mass in Fig. 4. This is due to the combination of the galaxy mass – metallicity relation and the natural increase over cosmic time of the mean mass of volume-limited galaxy samples in a  $\Lambda$ CDM cosmology.

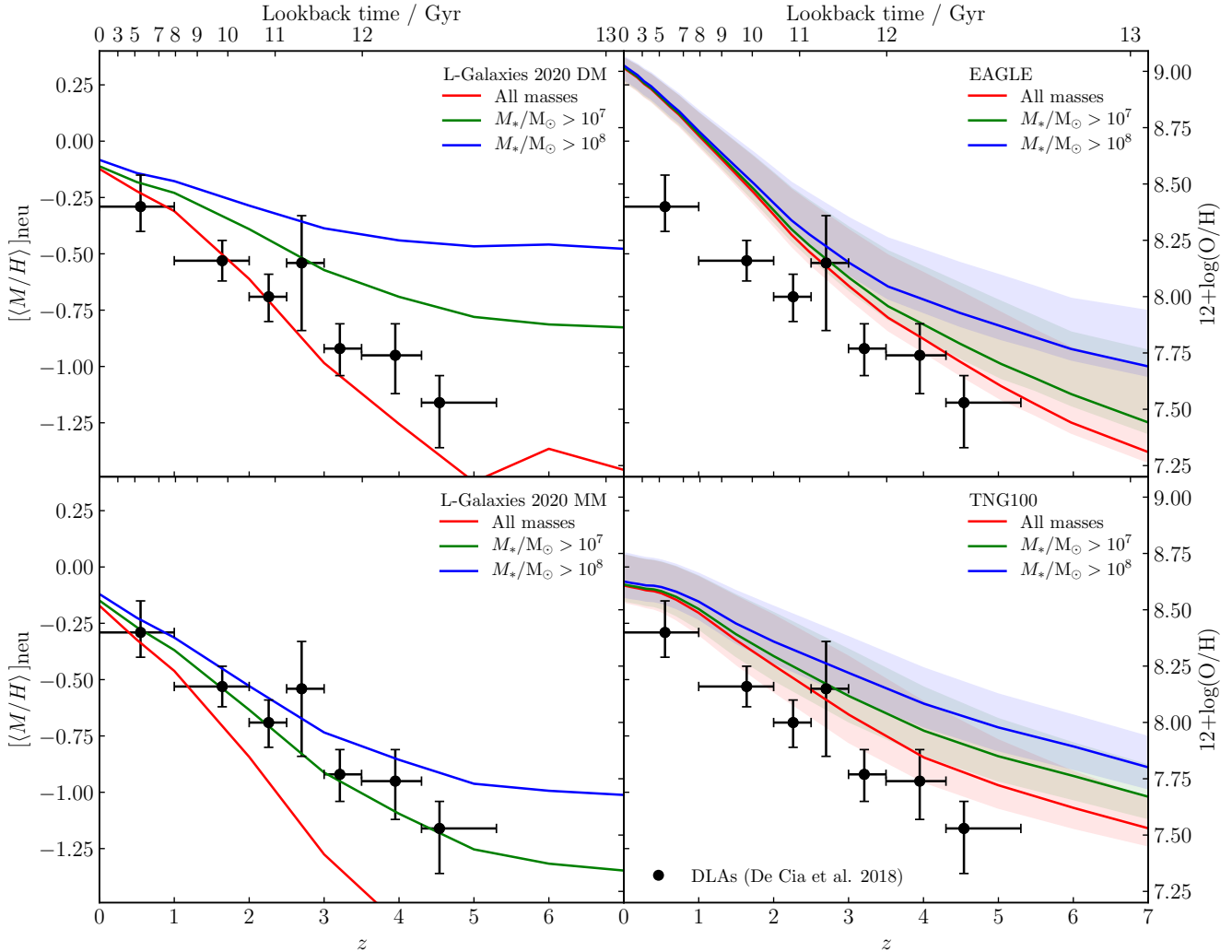
In L-GALAXIES 2020 DM, this evolution for all galaxies is close to that observed for DLAs, suggesting that current DLA samples might not have a significant mass bias compared to the overall galaxy population. However, we note again that L-GALAXIES 2020 DM produces overly-enriched galaxies at high redshift and is therefore disfavoured for this particular analysis. In L-GALAXIES 2020 MM, the evolution of  $[\langle M/H \rangle]_{\text{neu}}$  for all galaxies is *steeper* than observed for DLAs. Given that this simulation matches the observed metallicity evolution in ELGs at fixed mass (see Section 4.4.1), this suggests a shift in the mass distribution of DLA hosts with respect to the overall galaxy population at high redshift.

Dust corrections could also play a role, as these tend to be greater for metal-rich (i.e. lower-redshift) systems, thus influencing the slope of the observed  $[\langle M/H \rangle]_{\text{neu}}$  evolution for DLAs (see Krogager et al. 2019). However, the fact that the normalisation of  $[\langle M/H \rangle]_{\text{neu}}$  in L-GALAXIES 2020 MM is in good agreement with DLA observations at  $z \lesssim 1$  suggests that over-estimated dust depletion corrections in metal-rich systems are not a main factor here.

The most straightforward way to model a shift in the mass distribution of galaxies at high redshift in simulations is to impose a redshift-invariant lower-mass limit (see also Dvorkin et al. 2015). When doing this, Fig. 5 shows that the observed  $[\langle M/H \rangle]_{\text{neu}}$  evolution from DLAs is best reproduced in L-GALAXIES 2020 MM when only galaxies with masses above  $\log_{10}(M_*/M_{\odot}) \sim 7.0 - 7.5$  are considered.

We note here that galaxies of mass below  $\log_{10}(M_*/M_{\odot}) \sim 7.0$  live in DM subhalos with  $M_{200} \sim 4 \times 10^7 h^{-1} M_{\odot}$  in the simulations considered here, comprising only  $\sim 20$  DM particles. These systems are therefore at the resolution limit, and their baryonic properties are unlikely to be modelled accurately. Above  $\log_{10}(M_*/M_{\odot}) \sim 8.0$ , galaxies are hosted by subhaloes with  $M_{200} \gtrsim 1.3 \times 10^9 h^{-1} M_{\odot}$ , containing  $\gtrsim 100$  DM particles and are therefore substantially better resolved. The lower-mass limit for DLA host galaxies preferred by L-GALAXIES 2020 MM is therefore at the boundary between well and poorly-resolved galaxies in these simulations. Its value is therefore open to some uncertainty and we have not attempted to determine it more precisely here.

Nonetheless, a lower-mass limit of  $\log_{10}(M_*/M_{\odot}) \sim 7.5$  is consistent with recent determinations of DLA host masses in absorption-selected galaxy samples at  $z = 2 - 3$  (e.g. Krogager et al. 2020; Rhodin et al. 2021), although we note that such samples are biased towards metal-rich DLAs and therefore more massive systems. The lower limit determined from L-GALAXIES 2020 MM is also in good agreement with the approximate stellar mass estimates obtained using the metallicity- and redshift-dependent relation of Møller et al. (2013). When assuming a conversion factor between emission-line-based and absorption-line-based metallicities of  $C = 0.44 \pm 0.10$  from Chris-



**Figure 5.** Evolution of the mean, HI-weighted gas metallicity for DLAs (De Cia et al. 2018, black points) and the four simulations considered here. Solid lines represent the  $[M/H]_{\text{neu}}$  evolution for model galaxies above a certain mass threshold (see legend). Light shaded regions represent the uncertainty in the DLA density threshold for EAGLE and TNG100 (see Section 4.1.1). All metallicities are normalised to the solar photospheric value of  $M_{\text{met},\odot}/M_{\text{H},\odot} = 0.0181$  from Asplund et al. (2009).

tensen et al. (2014), we find that 97 per cent of the DLA hosts in the De Cia et al. (2018) sample are expected to have stellar masses above  $\log_{10}(M_*/M_{\odot}) = 7.5$ . The presence of a non-zero lower-mass limit and an increase in average mass over cosmic time is also consistent with the cross-correlation of DLAs and the Ly $\alpha$  forest (Pérez-Ràfols et al. 2018).

In L-GALAXIES 2020 MM, a lower-mass limit of  $\log_{10}(M_*/M_{\odot}) = 7.5$  returns an evolution in the *mean* stellar mass of galaxies from  $z = 5$  to 0 of 1.26 dex, increasing from  $\log_{10}(M_*/M_{\odot}) = 8.13$  to 9.38. Due to the shape of the galaxy stellar mass function, the equivalent evolution in the *median* stellar mass is only 0.18 dex, increasing from  $\log_{10}(M_*/M_{\odot}) = 7.76$  to 7.94. The mean virial mass of subhaloes increases by 0.61 dex from  $\log(M_{200}/M_{\odot}) = 10.57$  to 11.18 over the same period.

At  $z = 3$ , the mean virial mass of  $\log(M_{200}/M_{\odot}) = 10.74$  we find is in good agreement with the estimate of 10.78 determined by Krogager et al. (2020) from their DLA host galaxy modelling. Their model considers host galaxies down to  $\log_{10}(M_*/M_{\odot}) \sim 6.5$  at  $z = 2$ , which is an order of magnitude lower than the limit preferred

by L-GALAXIES 2020 MM. The good correspondence between the two models therefore further suggests that the lowest-mass systems do not contribute significantly to DLA samples, perhaps due to their lower gas cross-sections. Indeed, we find that at  $3 \lesssim z \lesssim 5$ , where the discrepancy between observational and model  $\rho_{\text{met,neu}}$  estimates is greatest, galaxies with  $\log_{10}(M_*/M_{\odot}) \gtrsim 7.5$  strongly dominate the  $\rho_{\text{met,neu}}$  budget in all four simulations, and sub-DLAs are expected to have only a minor contribution (e.g. Berg et al. 2021). Our findings are also consistent with those of Erkal et al. (2012), who used a high-resolution hydrodynamical simulation to determine that hosts with  $\log(M_{200}/M_{\odot}) < 11$  provide the dominant contribution to DLA cross sections at all column densities at  $z = 3$ .

Importantly, a lower-mass limit of  $\log_{10}(M_*/M_{\odot}) \sim 7.5$  also brings L-GALAXIES 2020 MM into better agreement with the observed  $\rho_{\text{met,neu}}$  from DLAs in Fig. 1 (dashed green line). However, an additional discrepancy still remains. This is likely due to the residual deficit in neutral gas present in subhaloes at high redshift in L-GALAXIES 2020, as discussed in Section 4.3.

EAGLE and TNG100 both return higher metallicities than seen

for DLAs at all redshifts, regardless of the lower-mass threshold assumed. This is particularly true for denser gas in EAGLE (upper limit of the uncertainty bands), and cannot be explained by differences among strong-emission-line metallicity diagnostics. This suggests that the higher  $\rho_{\text{met,neu}}$  at high redshift in EAGLE (see Fig. 1) could be partly due to an over-retention of metals in the dense ISM, at the expense of the partially-ionised gas phases. Whereas the higher  $\rho_{\text{met,neu}}$  seen in TNG100 may be due to an over-production of metals altogether (see Section 4.2).

#### 4.4.3 Other factors affecting $[\langle M/H \rangle]$ evolution

In addition to possible mass dependencies affecting the observed DLA metallicity evolution, a systematic evolution in ISM radial metallicity gradients could also play a role. However, observations of such gradients in galaxies back to  $z \sim 2.5$  suggest there is little to no overall evolution, with star-forming galaxies typically having slightly negative gradients at all redshifts (e.g. Stott et al. 2014; Leethochawalit et al. 2016; Patricio et al. 2019; Curti et al. 2020; Gillman et al. 2021). A similarly mild evolution in ISM metallicity gradients at fixed mass in galaxy discs is found in galaxy evolution simulations (e.g. Gibson et al. 2013; Ma et al. 2017; Yates et al. 2021), although Hemler et al. (2020) find a somewhat stronger evolution in TNG50.

A systematic evolution in typical DLA impact parameters,  $b$ , could also affect the inferred  $[\langle M/H \rangle]_{\text{neu}}$  evolution. This could be the case even without an evolution in typical ISM metallicity gradients, if DLAs also probe dense gas outside of galaxies. Indeed, Christensen et al. (2014) find a predicted decrease in the metallicity from the centre of host galaxies to the DLA itself of  $-0.022 \pm 0.004$  dex/kpc, with Krogager et al. (2020) similarly finding  $-0.019 \pm 0.008$  dex/kpc for  $L^*$  galaxies. However, the offset this implies between  $[\langle M/H \rangle]_{\text{neu}}$  measurements from DLAs and model galaxies is mitigated somewhat by the fact that the mean values calculated at each redshift are H $\alpha$  weighted, reducing the impact of DLAs with large  $b$ , as these systems tend to have lower H $\alpha$  column densities (see e.g. Krogager et al. 2017). In addition, we find that those few DLAs in the De Cia et al. (2018) sample which do have a large measured  $b$  of  $\geq 12$  kpc (Christensen et al. 2014; Augustin et al. 2018) all have higher metallicities than the average for DLAs at their redshift. Cold, dense gas clumps observed in galactic outflows are expected to be relatively metal rich (Péroux et al. 2020), although dense accreting gas should have lower metallicities (e.g. Garratt-Smithson et al. 2021).

Finally, systematic differences between emission-based and absorption-based metallicity measurements could also affect the comparisons made between various ISM metallicity measurements in this section. As mentioned in Section 2, emission-based abundances from QSO-DLA host galaxies can differ from the absorption-based measurements by up to  $\sim 0.6$  dex (Péroux et al. 2014; Rahmani et al. 2016). However, this difference includes an intrinsic contribution from any metallicity gradient present, as the QSO-DLAs used in those studies were offset from the central emission-line regions by 10 – 30 kpc. Systems with smaller impact parameters, such as SBS 1543+593 with  $b \sim 3$  kpc, instead return differences between absorption-based and emission-based metallicities of  $\leq 0.1$  dex (Schulte-Ladbeck et al. 2004, 2005; Bowen et al. 2005). Likewise, measurements for the host galaxy of GRB121024A, where the emission- and absorption-line regions are expected to be roughly co-spatial, suggest a discrepancy of only  $\sim 0.3$  dex (Friis et al. 2015). Studies of the blue compact dwarf galaxy, I Zw 18, which compare absorption-based abundances from its H $\alpha$  regions with *direct* emission-based abundances from electron temperature measure-

ments, also return a discrepancy of only  $\sim 0.17$  dex (Aloisi et al. 2003; Lecavelier des Etangs et al. 2004; Péquignot 2008; Lebouteiller et al. 2013). Nonetheless, further investigation, particularly at higher redshifts, is required to better determine any systematic differences between emission-line and absorption-line techniques.

In conclusion, we find that the mean H $\alpha$ -weighted metallicity in galaxies at high redshift in the L-GALAXIES 2020, EAGLE, and TNG100 simulations is sensitive to the minimum stellar mass of the population. Simulations can therefore be used in conjunction with the currently-observed  $[\langle M/H \rangle]_{\text{neu}}$  evolution to constrain the mass distribution of DLA host galaxies at high redshift. A lower-mass limit of  $\log_{10}(M_*/M_{\odot}) \sim 7.5$  would reconcile the L-GALAXIES 2020 MM with both observed DLA metallicities and  $\rho_{\text{met,neu}}$ . However, such an assumption does not improve the correspondence between observations and EAGLE or TNG100. These simulations over-predict DLA metallicities, possible due to an over-retention of metals in very dense gas in EAGLE, and an over-production of metals in TNG100. Other biases in observational samples could also play a role, and future studies of DLA host galaxies will be invaluable for improving constraints (see Section 5).

## 5 FUTURE PROSPECTS

As mentioned in Section 1, the dominant contributors to the total metal budget, other than DLAs, have not been measured at higher redshifts. The hot phase in particular, which is mainly traceable in absorption through e.g. Ne VIII in the UV (Frank et al. 2018; Burchett et al. 2019) and O VII and O VIII in the X-ray (Nelson et al. 2018b), are presently limited to  $z \lesssim 1.4$  (Ettori et al. 2015). Alternatives using novel techniques or metal transitions (e.g. Pettini & Dodorico 1986; Anderson & Sunyaev 2016, 2018; Fresco et al. 2020) will likely play a role. Ultimately, more robust measurements will have to await the next generation of X-ray facilities such as Athena (Nandra et al. 2013) and Lynx (The Lynx Team 2018).

At intermediate redshift, the census of metals in the Universe is still open. Ambitious surveys of order a million quasar spectra will provide a statistical sample of absorbers sufficient to obtain a definite picture of the metal content in baryons in the Universe over its entire lookback time. These efforts will rely on the availability of the next generation of multi-object spectrographs, e.g. 4MOST (de Jong et al. 2019), WEAVE (Dalton et al. 2012), and MSE (The MSE Science Team et al. 2019). Similarly, 21 cm measurements beyond  $z \sim 1$  from e.g. the MeerKAT LADUMA survey (Blyth et al. 2016) will help determine to what extent, if at all, DLAs provide biased  $\rho_{\text{HI}}$  estimates.

A more precise census of the star formation occurring in lower-mass galaxies at high redshift will also only be possible with the enhanced capabilities of JWST (Gardner et al. 2006). Instruments such as NIRCам and MIRI will allow SFRs to be measured for many faint galaxies at  $z \sim 2.5$  and beyond, as well as providing additional photometry in the rest-frame NIR for galaxies at  $z \geq 4$ , potentially helping to constrain stellar masses for high-redshift DLA hosts.

The unparalleled power of JWST/NIRSpec and VLT/BlueMUSE (Richard et al. 2019) will also enable a large number of precise gas-phase metallicity measurements in and around galaxies to be made at high redshift. These, along with detailed high-redshift metallicity calibrations from VLT/MOONS (Cirasuolo et al. 2020), will enable a much more detailed comparison between the metal density in DLA samples and that of the general star-forming galaxy population.

Finally, we highlight that the present study is concerned primarily with the cosmic metal density in the Universe, rather than the

metal content of DLAs specifically. The ideal way to study the latter in simulations would be to produce mock line-of-sight spectra and modelling a range of metal absorption lines from multiple metal ions such as Si, S, and Fe. Assuming the same parameters as in DLA observations (e.g. pixel resolution and signal-to-noise ratio) would also be important, because the column density of pixels along the line of sight and the damping wings are both needed to identify DLAs and compute their properties realistically. Some works have already adopted methods similar to this for selecting and analysing DLAs in simulations (e.g. Pontzen et al. 2008; Bird et al. 2014; Berry et al. 2016; Di Gioia et al. 2020; Hassan et al. 2020; Marra et al. 2021), and further effort in this direction would be invaluable as the focus of future work with galaxy evolution simulations.

## 6 CONCLUSIONS

In this work, we compare recent observational results on the cosmic density of metals, star formation, neutral gas, and metallicity with that found in four cosmological galaxy evolution simulations: EAGLE, TNG100, and two versions of L-GALAXIES 2020. This analysis allows us to investigate under what conditions the high  $\rho_{\text{met,neu}}$  estimates observed in high-redshift DLAs are possible from a theoretical perspective. Our main findings are as follows:

- The fraction of all metals that are in neutral gas at  $z \gtrsim 3$  is lower in models than inferred from DLAs. Observations of DLAs at  $z \sim 4$  suggest that  $\rho_{\text{met,neu}}$  constitutes  $\sim 87$  per cent of the total expected metal budget, when assuming the  $\rho_{\text{SFR}}$  from Madau & Dickinson (2014). In contrast, EAGLE, TNG100, and L-GALAXIES 2020 all suggest  $< 40$  per cent at the same redshift (see Section 4.1).
- In EAGLE and both versions of L-GALAXIES 2020, the majority of metals at all redshifts are found in hot, low-density gas, despite the range of feedback efficiencies used in these simulations. This phase is particularly difficult to observe at high redshift, and plays an important role in accounting for the ‘missing metals problem’. In TNG100, the dominant contribution at  $3 \lesssim z \lesssim 5$  is shared roughly equally between this low-density gas component, stars, and neutral gas (see Section 4.1).
- The absolute value of  $\rho_{\text{met,neu}}$  measured from DLAs at high redshift is well reproduced by TNG100. However, in order to achieve this, a larger  $\rho_{\text{SFR}}$  than is observed from FUV observations is needed. This tension highlights a possible incompatibility between the currently observed  $\rho_{\text{met,neu}}$  and  $\rho_{\text{SFR}}$  (see Section 4.2).
- The observed  $\rho_{\text{HI}}$  at  $z \gtrsim 3$  is significantly under-estimated by L-GALAXIES 2020, as has been seen in other semi-analytic models. This gas is most likely to be missing from the outskirts of resolved galaxies or the inner CGM, rather than from sub-resolution systems or the IGM. Accounting for this missing ISM material could help improve the correspondence between L-GALAXIES 2020 and the observed  $\rho_{\text{met,neu}}$  from DLAs (see Section 4.3).
- The discrepancy in  $\rho_{\text{met,neu}}$  between simulations and observations at high redshift could also be resolved if the mass distribution of DLA host galaxies is shifted to higher masses than the overall galaxy population. For example, when assuming a fixed lower-mass threshold of  $\log_{10}(M_*/M_{\odot}) \sim 7.5$ , the observed evolution in  $\rho_{\text{met,neu}}$ ,  $\rho_{\text{SFR}}$ , and  $[\langle M/H \rangle]_{\text{neu}}$  are simultaneously reproduced by L-GALAXIES 2020 MM (see Section 4.4).
- Both EAGLE and TNG100 appear to over-predict ISM metallicities for ELGs and DLAs at all redshifts. This could be partly due to an over-retention of metals in the densest gas in EAGLE, and an over-production of metals altogether in TNG100 (see Section 4.4),

although comparison to the redshift evolution of gas-phase metallicities is complicated by several systematic uncertainties.

We conclude that considering the simultaneous evolution of  $\rho_{\text{met,neu}}$ ,  $\rho_{\text{SFR}}$ ,  $\rho_{\text{HI}}$ , and  $[\langle M/H \rangle]_{\text{neu}}$  is crucial when trying to constrain the key physical processes driving galaxy evolution. Future observations of the metal content in hot, diffuse gas at high redshift, and of the host galaxy properties of DLAs, are also required to help break the degeneracies revealed between these key cosmic properties by cosmological galaxy formation simulations.

## ACKNOWLEDGEMENTS

The authors would like to thank the referee for their constructive and insightful comments, as well as Lise Christensen, Payel Das, Denis Erkal, Paolo Molaro, and Marta Spinelli for valuable discussions during the undertaking of this work. RMY acknowledges funding from the UK Research and Innovation council (grant number MR/S032223/1). CP thanks the Alexander von Humboldt Foundation for the granting of a Bessel Research Award held at MPA where this work was initiated. DN acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) through an Emmy Noether Research Group (grant number NE 2441/1-1). Part of this research was carried out using the High Performance Computing resources at the Max Planck Computing and Data Facility (MPCDF) in Garching, operated by the Max Planck Society (MPG).

## DATA AVAILABILITY

Data directly related to this publication and its figures is available at [www.tng-project.org/yates21](http://www.tng-project.org/yates21) and [rob Yatesastro.wixsite.com](http://rob Yatesastro.wixsite.com). The L-GALAXIES 2020 source code, as well as example output catalogues from the default and modified models, are publically available at [l galaxiespublicrelease.github.io/](https://l galaxiespublicrelease.github.io/). Complete output catalogues for the L-GALAXIES 2020 default model are also available on the MILLENNIUM database (Lemson & Virgo Consortium 2006) at [gavo.mpa-garching.mpg.de/Millennium/](http://gavo.mpa-garching.mpg.de/Millennium/). The IllustrisTNG simulations are publicly available and accessible in their entirety at [www.tng-project.org/data](http://www.tng-project.org/data) (Nelson et al. 2019a). The EAGLE simulation is publicly available (see McAlpine et al. 2016, for the original data release description).

## REFERENCES

- Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, *ApJ*, **584**, 45  
 Aloisi A., Savaglio S., Heckman T. M., Hoopes C. G., Leitherer C., Sembach K. R., 2003, *ApJ*, **595**, 760  
 Anderson M. E., Sunyaev R., 2016, *MNRAS*, **459**, 2806  
 Anderson M. E., Sunyaev R., 2018, *A&A*, **617**, A123  
 Anderson M. E., Bregman J. N., Butler S. C., Mullis C. R., 2009, *ApJ*, **698**, 317  
 Andrews B. H., Martini P., 2013, *ApJ*, **765**, 140  
 Angulo R. E., Hilbert S., 2015, *MNRAS*, **448**, 364  
 Angulo R. E., White S. D. M., 2010, *MNRAS*, **405**, 143  
 Asada Y., Ohta K., Maeda F., 2021, *ApJ*, **915**, 47  
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, **47**, 481  
 Augustin R., et al., 2018, *MNRAS*, **478**, 3120  
 Ayromlou M., Nelson D., Yates R. M., Kauffmann G., Renneby M., White S. D. M., 2021, *MNRAS*, **502**, 1051  
 Baldi A., Ettori S., Molendi S., Balestra I., Gastaldello F., Tozzi P., 2012, *A&A*, **537**, A142

- Balestra I., Tozzi P., Ettori S., Rosati P., Borgani S., Mainieri V., Norman C., Viola M., 2007, *A&A*, 462, 429
- Bellstedt S., et al., 2020, *MNRAS*, 498, 5581
- Berg M. A., et al., 2019, *ApJ*, 883, 5
- Berg T. A. M., et al., 2021, *MNRAS*, 502, 4009
- Berry M., Somerville R. S., Haas M. R., Gawiser E., Maller A., Popping G., Trager S. C., 2014, *MNRAS*, 441, 939
- Berry M., Somerville R. S., Gawiser E., Maller A. H., Popping G., Trager S. C., 2016, *MNRAS*, 458, 531
- Bird S., Vogelsberger M., Haehnelt M., Sijacki D., Genel S., Torrey P., Springel V., Hernquist L., 2014, *MNRAS*, 445, 2313
- Bischetti M., Maiolino R., Carniani S., Fiore F., Piconcelli E., Fluetsch A., 2019, *A&A*, 630, A59
- Blyth S., et al., 2016, in *MeerKAT Science: On the Pathway to the SKA*. p. 4
- Bouché N., Lehnert M. D., Péroux C., 2005, *MNRAS*, 364, 319
- Bouché N., Lehnert M. D., Péroux C., 2006, *MNRAS*, 367, L16
- Bouché N., Lehnert M. D., Aguirre A., Péroux C., Bergeron J., 2007, *MNRAS*, 378, 525
- Bowen D., Jenkins E., Pettini M., Tripp T., 2005, *ApJ*, 635, 880
- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150
- Burchett J. N., et al., 2019, *ApJ*, 877, L20
- Byrohl C., et al., 2021, *MNRAS*, 506, 5129
- Christensen L., Møller P., Fynbo J. P. U., Zafar T., 2014, *MNRAS*, 445, 225
- Cirasuolo M., et al., 2020, *The Messenger*, 180, 10
- Crain R. A., et al., 2015, *MNRAS*, 450, 1937
- Curti M., et al., 2020, *MNRAS*, 492, 821
- Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140
- Dalton G., et al., 2012, in *McLean I. S., Ramsay S. K., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV*. p. 84460P, doi:10.1117/12.925950
- De Cia A., 2018, *A&A*, 613, L2
- De Cia A., Ledoux C., Mattsson L., Petitjean P., Srianand R., Gavignaud I., Jenkins E. B., 2016, *A&A*, 596, A97
- De Cia A., Ledoux C., Petitjean P., Savaglio S., 2018, *A&A*, 611, A76
- De Rossi M. E., Bower R. G., Font A. S., Schaye J., Theuns T., 2017, *MNRAS*, 472, 3354
- Di Gioia S., Cristiani S., De Lucia G., Xie L., 2020, *MNRAS*, 497, 2469
- Diemer B., et al., 2019, *MNRAS*, 487, 1529
- Doherty C. L., Gil-Pons P., Lau H. H. B., Lattanzio J. C., Siess L., 2014, *MNRAS*, 437, 195
- Donnari M., et al., 2019, *MNRAS*, 485, 4817
- Dopita M. A., Kewley L. J., Sutherland R. S., Nicholls D. C., 2016, *Ap&SS*, 361, 61
- Driver S. P., et al., 2018, *MNRAS*, 475, 2891
- Dvorkin I., Silk J., Vangioni E., Petitjean P., Olive K. A., 2015, *MNRAS*, 452, L36
- Erkal D., Gnedin N. Y., Kravtsov A. V., 2012, *ApJ*, 761, 54
- Ettori S., Baldi A., Balestra I., Gastaldello F., Molendi S., Tozzi P., 2015, *A&A*, 578, A46
- Faucher-Giguère C.-A., 2020, *MNRAS*, 493, 1614
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Ferrara A., Scannapieco E., Bergeron J., 2005, *ApJ*, 634, L37
- Fishlock C. K., Karakas A. I., Lugaro M., Yong D., 2014, *ApJ*, 797, 44
- Frank S., Pieri M. M., Mathur S., Danforth C. W., Shull J. M., 2018, *MNRAS*, 476, 1356
- Fresco A. Y., Péroux C., Merloni A., Hamanowicz A., Szakacs R., 2020, *MNRAS*, 499, 5230
- Friis M., et al., 2015, *MNRAS*, 451, 167
- Fu J., Guo Q., Kauffmann G., Krumholz M. R., 2010, *MNRAS*, 409, 515
- Fu J., et al., 2013, *MNRAS*, 434, 1531
- Furlong M., et al., 2015, *MNRAS*, 450, 4486
- Fynbo J. P. U., et al., 2013, *MNRAS*, 436, 361
- Gardner J. P., et al., 2006, *Space Sci. Rev.*, 123, 485
- Garratt-Smithson L., Power C., Lagos C. d. P., Stevens A. R. H., Allison J. R., Sadler E. M., 2021, *MNRAS*, 501, 4396
- Genel S., et al., 2014, *MNRAS*, 445, 175
- Gibson B. K., Pilkington K., Brook C. B., Stinson G. S., Bailin J., 2013, *A&A*, 554, A47
- Gillman S., et al., 2021, *MNRAS*, 500, 4229
- Ginolfi M., et al., 2020, *A&A*, 633, A90
- Gnedin N. Y., Kravtsov A. V., 2011, *ApJ*, 728, 88
- Graham J. F., Schady P., Fruchter A. S., 2019, arXiv e-prints, p. arXiv:1904.02673
- Gruppioni C., et al., 2020, *A&A*, 643, A8
- Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, *MNRAS*, 428, 1351
- Haardt F., Madau P., 2001, in *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*. p. 64 (arXiv:astro-ph/0106018)
- Hassan S., Finlator K., Davé R., Churchill C. W., Prochaska J. X., 2020, *MNRAS*, 492, 2835
- Heintz K. E., Watson D., Oesch P., Narayanan D., Madden S. C., 2021, arXiv e-prints, p. arXiv:2108.13442
- Hemler Z. S., et al., 2020, arXiv e-prints, p. arXiv:2007.10993
- Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, *MNRAS*, 396, 535
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663
- Henriques B. M. B., Yates R. M., Fu J., Guo Q., Kauffmann G., Srisawat C., Thomas P. A., White S. D. M., 2020, *MNRAS*, 491, 5795
- Hunt L., Dayal P., Magrini L., Ferrara A., 2016, *MNRAS*, 463, 2002
- Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., 2015, *MNRAS*, 451, 2251
- Jenkins E. B., 2009, *ApJ*, 700, 1299
- Jenkins E. B., Wallerstein G., 2017, *ApJ*, 838, 85
- Karakas A. I., 2010, *MNRAS*, 403, 1413
- Katsianis A., et al., 2017, *MNRAS*, 472, 919
- Katsianis A., et al., 2020, *MNRAS*, 492, 5592
- Kennicutt Robert C. J., 1998, *ApJ*, 498, 541
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kistler M. D., Yüksel H., Beacom J. F., Hopkins A. M., Wyithe J. S. B., 2009, *ApJ*, 705, L104
- Kobayashi C., Umeda H., Nomoto K., Tominaga N., Ohkubo T., 2006, *ApJ*, 653, 1145
- Kobulnicky H. A., Kewley L. J., 2004, *ApJ*, 617, 240
- Koushan S., et al., 2021, *MNRAS*, 503, 2033
- Krogager J., et al., 2013, *MNRAS*, 433, 3091
- Krogager J.-K., Møller P., Fynbo J. P. U., Noterdaeme P., 2017, *MNRAS*, 469, 2959
- Krogager J.-K., Fynbo J. P. U., Møller P., Noterdaeme P., Heintz K. E., Pettini M., 2019, *MNRAS*, 486, 4377
- Krogager J.-K., Møller P., Christensen L. B., Noterdaeme P., Fynbo J. P. U., Freudling W., 2020, *MNRAS*, 495, 3014
- Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, *MNRAS*, 418, 1649
- Lagos C. D. P., Baugh C. M., Zwaan M. A., Lacey C. G., Gonzalez-Perez V., Power C., Swinbank A. M., van Kampen E., 2014, *MNRAS*, 440, 920
- Lebouteiller V., Heap S., Hubeny I., Kunth D., 2013, *A&A*, 553, 16
- Lecavelier des Etangs A., Désert J. M., Kunth D., Vidal-Madjar A., Callejo G., Ferlet R., Hébrard G., Lebouteiller V., 2004, *A&A*, 413, 131
- Leethochawalit N., Jones T. A., Ellis R. S., Stark D. P., Richard J., Zitrin A., Auger M., 2016, *ApJ*, 820, 84
- Leja J., et al., 2019, *ApJ*, 877, 140
- Leja J., Speagle J. S., Johnson B. D., Conroy C., van Dokkum P., Franx M., 2020, *ApJ*, 893, 111
- Lemson G., Virgo Consortium t., 2006, arXiv e-prints, pp astro-ph/0608019
- Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017, *MNRAS*, 466, 4780
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Maiolino R., et al., 2008, *A&A*, 488, 463
- Marigo P., 2001, *A&A*, 370, 194
- Marinacci F., et al., 2018, *MNRAS*, 480, 5113
- Marra R., et al., 2021, arXiv e-prints, p. arXiv:2108.03254
- Martindale H., Thomas P. A., Henriques B. M., Loveday J., 2017, *MNRAS*, 472, 1981

- Matthews A. M., Condon J. J., Cotton W. D., Mauch T., 2021, *ApJ*, **914**, 126
- McAlpine S., et al., 2016, *Astronomy and Computing*, **15**, 72
- McDonald M., et al., 2016, *ApJ*, **826**, L24
- McKee C. F., Krumholz M. R., 2010, *ApJ*, **709**, 308
- Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020, *MNRAS*, **494**, 3971
- Møller P., Fynbo J., Ledoux C., Nilsson K., 2013, *MNRAS*, **430**, 2680
- Mushotzky R. F., Loewenstein M., 1997, *ApJ*, **481**, L63
- Naiman J. P., et al., 2018, *MNRAS*, **477**, 1206
- Nandra K., et al., 2013, arXiv e-prints, p. [arXiv:1306.2307](https://arxiv.org/abs/1306.2307)
- Nelson D., et al., 2018a, *MNRAS*, **475**, 624
- Nelson D., et al., 2018b, *MNRAS*, **477**, 450
- Nelson D., et al., 2019a, *Computational Astrophysics and Cosmology*, **6**, 2
- Nelson D., et al., 2019b, *MNRAS*, **490**, 3234
- Nelson E. J., et al., 2021, *MNRAS*, **508**, 219
- Nomoto K., Hashimoto M., Tsujimoto T., Thielemann F. K., Kishimoto N., Kubo Y., Nakasato N., 1997, *Nuclear Phys. A*, **616**, 79
- Pagel B. E. J., 1999, arXiv e-prints, pp [astro-ph/9911204](https://arxiv.org/abs/astro-ph/9911204)
- Patrício V., et al., 2019, *MNRAS*, **489**, 224
- Peeples M. S., et al., 2019, *ApJ*, **873**, 129
- Péquignot D., 2008, *A&A*, **478**, 371
- Pérez-Ráfols I., et al., 2018, *MNRAS*, **473**, 3019
- Péroux C., Howk J. C., 2020, *ARA&A*, **58**, 363
- Péroux C., Kulkarni V. P., York D. G., 2014, *MNRAS*, **437**, 3144
- Péroux C., Nelson D., van de Voort F., Pillepich A., Marinacci F., Vogelsberger M., Hernquist L., 2020, *MNRAS*, **499**, 2462
- Pettini M., 1999, in Walsh J. R., Rosa M. R., eds, *Chemical Evolution from Zero to High Redshift*. p. 233 ([arXiv:astro-ph/9902173](https://arxiv.org/abs/astro-ph/9902173))
- Pettini M., Dodorico S., 1986, *ApJ*, **310**, 700
- Pettini M., Pagel B. E. J., 2004, *MNRAS*, **348**, L59
- Pillepich A., et al., 2018a, *MNRAS*, **473**, 4077
- Pillepich A., et al., 2018b, *MNRAS*, **475**, 648
- Pillepich A., et al., 2019, *MNRAS*, **490**, 3196
- Planck Collaboration 2016, *A&A*, **594**, A13
- Planck Collaboration et al., 2014a, *A&A*, **571**, A1
- Planck Collaboration et al., 2014b, *A&A*, **571**, A16
- Pontzen A., et al., 2008, *MNRAS*, **390**, 1349
- Popping G., et al., 2019, *ApJ*, **882**, 137
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, **334**, 505
- Rafelski M., Neeleman M., Fumagalli M., Wolfe A. M., Prochaska J. X., 2014, *ApJ*, **782**, L29
- Rahmani H., et al., 2016, *MNRAS*, **463**, 980
- Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, *MNRAS*, **430**, 2427
- Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, *MNRAS*, **452**, 2034
- Renzini A., 1998, arXiv e-prints, pp [astro-ph/9810304](https://arxiv.org/abs/astro-ph/9810304)
- Rhodin N. H. P., Krogager J. K., Christensen L., Valentino F., Heintz K. E., Møller P., Zafar T., Fynbo J. P. U., 2021, *MNRAS*, **506**, 546
- Richard J., et al., 2019, arXiv e-prints, p. [arXiv:1906.01657](https://arxiv.org/abs/1906.01657)
- Rowan-Robinson M., et al., 2016, *MNRAS*, **461**, 1100
- Sanders R. L., et al., 2016, *ApJ*, **825**, L23
- Sanders R. L., et al., 2021, *ApJ*, **914**, 19
- Schaye J., et al., 2010, *MNRAS*, **402**, 1536
- Schaye J., et al., 2015, *MNRAS*, **446**, 521
- Schulte-Ladbeck R., Rao S., Drozdovsky I., Turnshek D., Pettini M., 2004, *ApJ*, **600**, 613
- Schulte-Ladbeck R., König B., Miller C., Hopkins A., Drozdovsky I., Turnshek D., Hopp U., 2005, *ApJ*, **625**, L79
- Shull J. M., Danforth C. W., Tilton E. M., 2014, *ApJ*, **796**, 49
- Spilker J. S., et al., 2020, *ApJ*, **905**, 86
- Spinelli M., Zoldan A., De Lucia G., Xie L., Viel M., 2020, *MNRAS*, **493**, 5434
- Springel V., 2010, *MNRAS*, **401**, 791
- Springel V., Hernquist L., 2003, *MNRAS*, **339**, 289
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, **328**, 726
- Springel V., Di Matteo T., Hernquist L., 2005, *MNRAS*, **361**, 776
- Springel V., et al., 2018, *MNRAS*, **475**, 676
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, **717**, 289
- Stevens A. R. H., et al., 2019, *MNRAS*, **483**, 5334
- Stott J. P., et al., 2014, *MNRAS*, **443**, 2695
- Sugahara Y., Ouchi M., Harikane Y., Bouché N., Mitchell P. D., Blaizot J., 2019, *ApJ*, **886**, 29
- The Lynx Team 2018, arXiv e-prints, p. [arXiv:1809.09642](https://arxiv.org/abs/1809.09642)
- The MSE Science Team et al., 2019, arXiv e-prints, p. [arXiv:1904.04907](https://arxiv.org/abs/1904.04907)
- Theuns T., 2021, *MNRAS*, **500**, 2741
- Thielemann F. K., et al., 2003, in Hillebrandt W., Leibundgut B., eds, *From Twilight to Highlight: The Physics of Supernovae*. p. 331, doi:10.1007/10828549\_46
- Thöne C. C., et al., 2013, *MNRAS*, **428**, 3590
- Torrey P., et al., 2019, *MNRAS*, **484**, 5587
- Tozzi P., Rosati P., Etori S., Borgani S., Mainieri V., Norman C., 2003, *ApJ*, **593**, 705
- Tremonti C. A., et al., 2004, *ApJ*, **613**, 898
- Utomo D., Kriek M., Labbé I., Conroy C., Fumagalli M., 2014, *ApJ*, **783**, L30
- Villaescusa-Navarro F., et al., 2018, *ApJ*, **866**, 135
- Vladilo G., Bonifacio P., Centurión M., Molaro P., 2000, *ApJ*, **543**, 24
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, **436**, 3031
- Wang L., Pearson W. J., Cowley W., Trayford J. W., Béthermin M., Gruppioni C., Hurley P., Michałowski M. J., 2019, *A&A*, **624**, A98
- Weinberger R., et al., 2017, *MNRAS*, **465**, 3291
- Wiersma R. P. C., Schaye J., Smith B. D., 2009, *MNRAS*, **393**, 99
- Wilks S. M., Lovell C. C., Stanway E. R., 2019, *MNRAS*, **490**, 5359
- Wolfe A. M., Gawiser E., Prochaska J. X., 2005, *ARA&A*, **43**, 861
- Xie L., De Lucia G., Hirschmann M., Fontanot F., Zoldan A., 2017, *MNRAS*, **469**, 968
- Yates R. M., Henriques B., Thomas P. A., Kauffmann G., Johansson J., White S. D. M., 2013, *MNRAS*, **435**, 3500
- Yates R. M., Schady P., Chen T. W., Schweyer T., Wiseman P., 2020, *A&A*, **634**, A107
- Yates R. M., Henriques B. M. B., Fu J., Kauffmann G., Thomas P. A., Guo Q., White S. D. M., Schady P., 2021, *MNRAS*, **503**, 4474
- Zafar T., Péroux C., Popping A., Milliard B., Deharveng J. M., Frank S., 2013, *A&A*, **556**, A141
- Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014, *ApJ*, **791**, 130
- Zavala J. A., et al., 2021, *ApJ*, **909**, 165
- Zhao P., Xu H., Katsianis A., Yang X.-H., 2020, *Research in Astronomy and Astrophysics*, **20**, 195
- Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, *MNRAS*, **359**, L30
- de Jong R. S., et al., 2019, *The Messenger*, **175**, 3