Interplay of Stellar and Gas-Phase Metallicities: Unveiling Insights for Stellar Feedback Modeling with Illustris, IllustrisTNG, and EAGLE

Alex M. Garcia[1,2]⋆, Paul Torrey[1,2], Kathryn Grasha[3,5,6]†, Lars Hernquist[4], Sara Ellison[7], Henry R.M. Zovaro[3,5], Z.S. Hemler[8], Erica J. Nelson[9], Lisa J. Kewley[4]

[1]*Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA*

[2]*Department of Astronomy, University of Florida, 211 Bryant Space Sciences Center, Gainesville, FL 32611, USA*

[3]*Research School of Astronomy & Astrophysics, Australian National University, Canberra, Australia, 2611*

[4]*Institute for Theory and Computation, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA*

[5]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*

[6]*Visiting Fellow, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

[7]*Department of Physics & Astronomy, University of Victoria, Finnerty Road, Victoria, British Columbia, V8P 1A1, Canada*

[8]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ, 08544, USA*

[9]*Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80309, USA*

ABSTRACT

The metal content of galaxies provides a window into their formation in the full context of the cosmic baryon cycle. In this study, we examine the relationship between stellar mass and stellar metallicity (MZ$_*$R) in the hydrodynamic simulations Illustris, TNG, and EAGLE to understand the global properties of stellar metallicities within the feedback paradigm employed by these simulations. Interestingly, we observe significant variations in the overall normalization and redshift evolution of the MZ$_*$R across the three simulations. However, all simulations consistently demonstrate a tertiary dependence on the specific star formation rate (sSFR) of galaxies. This finding parallels the relationship seen in both simulations and observations between stellar mass, gas-phase metallicity, and some proxy of galaxy gas content (e.g., SFR, gas fraction, atomic gas mass). Since we find this correlation exists in all three simulations, each employing a sub-grid treatment of the dense, star-forming interstellar medium (ISM) to simulate smooth stellar feedback, we interpret this result as a fairly general feature of simulations of this kind. Furthermore, with a toy analytic model, we propose that the tertiary correlation in the stellar component is sensitive to the extent of the "burstiness" of feedback within galaxies.

Key words: galaxies: abundances – galaxies: evolution – galaxies: ISM – ISM: abundances

1 INTRODUCTION

Heavy elements are valuable tracers of the gas flows that are an integral part of galaxy formation. As aging stellar populations return metals to the interstellar medium (ISM), these metals are then dispersed and advected with any subsequent gas flows, and are incorporated into future generations of stars (Kobayashi et al. 2020). Thus, the metal content of galaxies is a balance that is sensitive not only to the amount of heavy elements produced, but also on the efficacy of gas mixing within the ISM (e.g., Elmegreen 1999; Veilleux et al. 2005), prevalence of outflows (e.g., Veilleux et al. 2020), and rate of fresh/pristine gas accretion (Kereš et al. 2005). Since the metals are intrinsically tied to these physical processes within a galaxy, they may act as an observable diagnostic for the physics driving galaxy evolution (Dalcanton 2007; Kewley et al. 2019). For example, since outflows depend strongly on stellar feedback within a galaxy and out-

flows have a distinct impact on galactic metallicity, the metal content of galaxies can thus be used as a probe of stellar feedback.

On large scales, trends between physical properties and metallicity within galaxy populations have been observed. For example, the stellar mass-gas-phase metallicity relation (MZ$_g$R) describes a trend of increasing metallicity with increasing galaxy mass (Tremonti et al. 2004; Lee et al. 2006) and is regarded as a key scaling relation between the stellar masses of galaxies and the metallicity of the gas component. The reason that high mass galaxies correspond to elevated metallicities has been postulated to be a result of the high mass galaxies having either higher effective yields (i.e. retaining a higher fraction of their produced metals in the gas-phase; Tremonti et al. 2004), lower gas fractions (which results in higher metallicities at a fixed amount of metal mass; Pasquali et al. 2012), or perhaps both. The slope of the MZ$_g$R has been noted to change as a function of mass with a steep power-law relation at low and intermediate masses that flattens at the highest masses (e.g., Blanc et al. 2019). These changes in slope are attributed to the efficiency (or lack thereof) of a number of different physical processes, including (but not limited to) stellar feedback, ISM turbulence, and the mixing of enriched and

pristine gas in the circumgalactic medium (CGM). Moreover, there is a marked evolution in the MZ$_g$R with time, whereby higher redshift galaxies form a clear MZ$_g$R but with a lower overall normalisation (e.g., Savaglio et al. 2005; Maiolino et al. 2008; Zahid et al. 2011; Langeroodi et al. 2022). The redshift evolution is attributed to the lack of chemical maturity at high redshift wherein fewer generations of stars have formed to synthesize heavier elements. The MZ$_g$R itself helps to reveal the net trends that must be present in terms of the gas mass evolution – coupled to the outflow properties – with galaxy mass and redshift for the metal content of galaxies to appropriately evolve.

In addition to the MZ$_g$R, there is also a stellar mass-stellar metallicity relation (MZ$_*$R) whereby the stellar metallicity increases with increasing stellar mass much in the same way as the MZ$_g$R (e.g. Gallazzi et al. 2005; Panter et al. 2008). This relationship, much like its gas-phase counterpart, follows a trend of increasing metals with increasing stellar mass before flattening at high masses (e.g., Sommariva et al. 2012; Zahid et al. 2017). Just as the overall normalisation of the MZ$_g$R decreases as a function of redshift, so too does the normalisation of the MZ$_*$R (e.g., Gallazzi et al. 2014; Cullen et al. 2019).

Many studies tend to focus on just the gas-phase or stellar metallicities, yet recent works (Topping et al. 2020; Cullen et al. 2021; Fraser-McKelvie et al. 2022; Greener et al. 2022, etc) have shown that there is potentially a large amount of power in examining them in conjunction. Physically, the link between stellar and gas-phase metals follows from the idea that newly formed stellar populations simply inherit the metallicity from the gas from which they were formed. Thus, the stellar metallicity of the galaxy reflects the mass (or luminosity) weighted average metallicity integrated over the galaxy's formation history. Conversely, the gas-phase traces recent evolutionary processes. As such, the stellar metallicity is not directly influenced by gas inflows and/or metal production from aging stellar populations, but is instead constantly producing new stellar populations that reflect the gas phase metallicity value at the time of formation. Stellar metallicities tend to be systematically lower than their gas-phase counterparts across stellar mass (e.g., Gallazzi et al. 2005; Lian et al. 2018). In more detail, González Delgado et al. (2014) find that the largest divergence of the two metallicities is in the lowest mass galaxies. Fraser-McKelvie et al. (2022) attribute this divergence to the efficacy of inflows and outflows at different masses as well as a strong dependence on a galaxy's star formation history (Lian et al. 2018 indicate that initial mass function variations are another possibility). It should be noted that the absolute values of gas-phase metallicites are sensitive to metallicity diagnostic (Kewley & Ellison 2008). Therefore trends between gas-phase and stellar metallicities are likely also sensitive to the chosen diagnostic. The magnitude of the discrepancy between the two metallicities in simulations has been seen to be both larger (Yates et al. 2012) and smaller (De Rossi et al. 2017) than observations. Comparisons between simulated metallicities and observed metallicities is not always straightforward, however, without the use of mock observational techniques (i.e., radiative transfer codes to generate spectra; see discussion in Nelson et al. 2018).

Very high redshift ($z \sim 8$) measurements of gas-phase metallicities have already been attained (e.g., Langeroodi et al. 2022; Shapley et al. 2023; Sanders et al. 2023), but stellar metallicities at these high redshifts are significantly more difficult to measure. The difference between the viability of obtaining gas-phase and stellar metallicities at high z can be attributed to the different methods of deriving the two metallicities. Briefly, gas-phase metallicities are derived from emission diagnostics, more specifically the direct and strong line methods (see Kewley & Ellison 2008; Kewley et al. 2019, for com-

plete reviews). Stellar metallicities, on the other hand, are typically determined by absorption features in the stellar continuum. The stellar diagnostics require high continuum signal-to-noise and as such, at present, the majority of stellar metallicity measurements come from low redshift systems (e.g., Chisholm et al. 2019), stacking galaxies (e.g., Halliday et al. 2008; Cullen et al. 2019; Sextl et al. 2023), observing lensed galaxies (e.g., Patrício et al. 2016), or very long exposure times (Kriek et al. 2019).

It is now recognized that the scatter of galaxies about the mass-metallicity relations is not random. Instead, systems with metallicities above (below) the MZ$_g$R tend to have low (high) gas fractions and low (high) star formation rates (Ellison et al. 2008; Mannucci et al. 2010; Lara-López et al. 2010; Bothwell et al. 2013, 2016; Yang et al. 2022, etc). This "correlated scatter" is naturally driven by the relationship between gas inflows, star formation, and metal production (though this is only true on the global scale; Baker et al. 2023). More specifically, gas inflows tend to drive down galactic metallicities while driving up gas content and star formation rates. Conversely, galaxies with low gas inflow rates steadily consume their gas, driving up metallicities while the gas fractions and star formation rates decay (Torrey et al. 2018; hereafter T18). This correlated scatter can therefore be thought of as a three-parameter relation between stellar mass, gas-phase metallicity, and some proxy for the gas richness of the galaxy (gas mass, SFR, etc.). Perhaps even more remarkably, the same three parameter fit that relates stellar mass, gas-phase metallicity, and gas richness can simultaneously fit galaxies at low- and high-redshift (Cresci et al. 2019; Huang et al. 2019, however recent JWST observations suggest that galaxies at $z \gtrsim 6$ deviate from the three parameter fit of lower redshift systems, see Curti et al. 2023). This so-called Fundamental Metallicity Relation (FMR) hints that the redshift evolution of the MZ$_g$R is not disjoint from its own scatter. Instead, the same physical processes (e.g., inflows, gas consumption) that drive galaxies about the MZ$_g$R at any given redshift, also cause an evolution in the MZ$_g$R with redshift.

However, since the stellar metallicities, and therefore the MZ$_*$R, are not directly influenced by gas inflows, it is not immediately obvious that there should be an equivalent FMR for stars. Yet, observationally (e.g., Zahid et al. 2017; Sextl et al. 2023) and in simulations (De Rossi et al. 2018), there is evidence for this, insofar as a relationship between the MZ$_*$R and star formation exists. As of yet, there have been no observational studies systematically studying or quantifying the existence of an "FMR for stars". In simulations, however, there is some evidence for this relationship not being "fundamental" as redshift evolution has been seen (Fontanot et al. 2021).

However, within simulations as a whole, significant uncertainty still exists in the preferred method to model both star formation and outflows in galaxy formation models – on which the metallicity critically depends. While empirical relations such as the Kennicutt-Schmidt relation (Schmidt 1959; Kennicutt 1998; KS relation) provide guidance or constraints for setting star formation rates (SFRs) based on the current gas content within galaxies, the implications for galactic star formation rates in very low mass or high redshift galaxies remains unclear. Indeed, many cosmological hydrodynamic simulations (e.g., Illustris, TNG, EAGLE) employ sub-grid ISM pressurization that manifestly enforce the KS relation (Vogelsberger et al. 2014b; Pillepich et al. 2018b; Crain et al. 2015), giving rise to SFRs that evolve reasonably smoothly out to the highest redshifts. Yet, models that attempt to treat both the multi-phase nature of the ISM and locality of star formation and feedback generally predict SFRs more explicitly with a significantly increased level of time variability. (i.e., burstiness; Faucher-Giguère 2018).

Notably, both "smooth" and "bursty" galaxy formation models are

	Illustris	TNG	EAGLE
Box side-length [Mpc]	106.5	110.7	100
Resolution Elements	2×1820^3	2×1820^3	2×1504^3
Sub-Grid Physics Model	SH03	SH03	SDV08
m_{baryon} [$10^6 M_\odot$]	1.26	1.4	1.81

Table 1. Summary of relevant properties of the three simulation models used in this work (Illustris, TNG, and EAGLE). m_{baryon} is the initial baryonic matter mass resolution in each simulation.

able to match a wide range of galaxy scaling relations (e.g., the galaxy stellar mass function, star formation main sequence, mass-metallicity relation; see, e.g., Hopkins et al. 2014; Vogelsberger et al. 2014a). Yet, the burstiness or time-variability of star formation (and its associated feedback) is of somewhat critical importance to our understanding of galaxy formation. In particular, bursty feedback – and the rapid, periodic, expulsion of material from the galactic center – has been invoked as a preferred mechanism that converts dark matter cuspy profiles into cored profiles – alleviating an important tension between ΛCDM and observations (e.g., Lazar et al. 2020).

In this paper, we examine this relationship between the MZ$_*$R and star formation using the Illustris, IllustrisTNG, and EAGLE hydrodynamical galaxy formation simulations. All three of these simulations are sufficiently large to capture galaxy populations where not only can the MZ$_*$R be defined, but the scatter can also be characterized. Notably, while these simulations share several similarities in how the ISM, star formation, and metal enrichment are treated, they also model the feedback processes that shape both the MZ$_g$R and MZ$_*$R in detail differently. As such, these simulations can be used to gain insight into variations resulting from changes in adopted galaxy formation models. Moreover, we develop an analytical model to probe not only how and why the residual correlation emerges, but also how its properties change with redshift.

The structure of this paper is as follows. In Section 2 we describe our methods including a brief description of the simulations employed, and our definitions of stellar metallicity. In Section 3 we present our results including a comparison of the MZ$_*$R's in Illustris, IllustrisTNG, and EAGLE, the residual correlations in the scatter, and its origin. In Section 4 we discuss our results, including a presentation of a simple toy model that can be used to capture the essential properties of the residual correlations, including the redshift evolution. Finally, in Section 5 we provide conclusions.

2 METHODS

In this work, we analyse the gas-phase and stellar metallicities in Illustris, IllustrisTNG, and EAGLE galaxies to understand their dependence on star formation and stellar mass (i.e., fundamental metallicity relations). Importantly, as will be detailed in the following sections, these three models have a commonality: sub-grid ISM pressurisation and smooth stellar feedback. Thus, though the detailed implementations of all the models are appreciably different, we believe that any results found in all of the models should comprise a fairly generic result of simulations with this ISM treatment; however, more testing is required within (and without, as we discuss in Section 4.2.1) the sub-grid ISM paradigm to confirm this.

In this section, we briefly describe each of the simulations and our selection criteria for galaxies within the sample analysed. All measurements are in physical units, except for the simulation box

sizes, which are reported in co-moving units. We summarize the details of this section in Table 1.

2.1 Simulation Details

2.1.1 Illustris

For our analysis, we utilize the Illustris (Vogelsberger et al. 2013, 2014a,b; Genel et al. 2014; Torrey et al. 2014) cosmological simulations. These runs were performed with the moving-mesh hydrodynamical code AREPO (Springel 2010). The Illustris model includes several important astrophysical processes including gravity, hydrodynamics, star formation, stellar evolution, chemical enrichment, radiative cooling and heating of the ISM, stellar feedback, black hole growth, and active galactic nuclei (AGN) feedback. Most important for the purposes of our work are self consistent implementations of star formation, stellar feedback, and chemical enrichment. Illustris treats the dense, star forming ISM with the effective equation of state of Springel & Hernquist (2003; hereafter, SH03). In this model, gas beyond a threshold density ($n_{\mathrm{H}} > 0.13 \ \mathrm{cm}^{-3}$) forms stars probabilistically assuming a Chabrier (2003) initial mass function (IMF). Crucially, the stars that are formed adopt their metallicity from the gas from which they form. As the stars evolve, they return both mass and metals to the ISM. Stellar mass return and yield tables are employed to allow for the direct simulation of time-dependent mass return and heavy element enrichment. Illustris explicitly tracks nine different chemical species (H, He, C, N, O, Ne, Mg, Si, and Fe).

The main Illustris simulation included a single volume of size $(106.5 \ \mathrm{Mpc})^3$ at three different resolutions, which are denoted as: Illustris-1 for the highest resolution run (with 2×1820^3 particles) and Illustris-3 for the lowest resolution run (with 2×455^3 particles). In this work, we use the highest resolution run, Illustris-1, which we will henceforth use synonymously with Illustris itself.

2.1.2 IllustrisTNG

IllustrisTNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018b; Springel et al. 2018; Pillepich et al. 2019; Nelson et al. 2019a,b, hereafter TNG) is the successor to the original Illustris simulations. TNG includes updates to the original Illustris model in addition to alleviating some deficiencies. Importantly, TNG also models the dense, star forming ISM with the SH03 equation of state. As such, these two simulations are related, yet have appreciably different physical implementations (for a complete list of differences see Weinberger et al. 2017; Pillepich et al. 2018a). Critical for the discussion later in this work, TNG implements redshift dependent winds. A wind velocity floor is enforced in TNG to ensure that low mass haloes do not have unphysically large mass loading factors. The effect of this is that winds are more efficient at suppressing low-redshift star formation and, consequently, feedback is more effective at higher redshifts. TNG tracks the same nine elements as Illustris, but it additionally adds an "other metals" item as a proxy for other metals not explicitly tracked.

Whereas the original Illustris suite consisted of only one volume at several resolution levels, TNG is comprised of three different volumes, each with their own set of resolutions. The simulations within TNG are designated by their approximate box size – TNG50 $(51.7 \ \mathrm{Mpc})^3$, TNG100 $(110.7 \ \mathrm{Mpc})^3$, and TNG300 $(302.6 \ \mathrm{Mpc})^3$, with the different resolution levels indicated as in the Illustris suite. Here, we employ the highest resolution run of TNG100 (TNG100-1; 2×1820^3 particles, hereafter simply TNG) to make a fair comparison with the original Illustris simulation.

2.1.3 EAGLE

Finally, we use the "Evolution and Assembly of GaLaxies and their Environment" (EAGLE, Crain et al. 2015; Schaye et al. 2015; McAlpine et al. 2016) cosmological simulations. Unlike Illustris and TNG, EAGLE employs a heavily modified version of GADGET-3 (Springel 2005) employing the ANARCHY formulation of smoothed particle hydrodynamics (SPH; see Schaye et al. 2015 Appendix A), which alleviates problems with the original version of SPH (Sijacki et al. 2012; Vogelsberger et al. 2012; Kereš et al. 2012; Torrey et al. 2012). EAGLE is a full-physics simulation that includes many of the same baryonic processes with hydrodynamics (star-formation, chemical enrichment, radiative cooling and heating, etc). The dense (unresolved) ISM is treated with a sub-grid equation of state (Schaye & Dalla Vecchia 2008; hereafter, SDV08), much like that of SH03. The SDV08 prescription forms stars according to a Chabrier (2003) IMF from the dense ISM gas. The density threshold for star formation is given by the metallicity-dependent transition from atomic to molecular gas computed by Schaye (2004), with an additional temperature-dependent criterion (see Schaye et al. 2015, their Section 4.3). Stellar populations evolve according to the Wiersma et al. (2009) evolutionary model and eventually return their mass and metals back into the ISM. While the SH03 equation of state describes the multiphase ISM with the average gas density, SDV08 use a polytropic equation of state. EAGLE tracks eleven different chemical species (H, He, C, N, O, Ne, Mg, Si, S, Ca, and Fe).

EAGLE is comprised of several simulations ranging from size $(12\,\mathrm{Mpc})^3$ to $(100\,\mathrm{Mpc})^3$. For this work, as an even-handed comparison to the selected Illustris and TNG runs, we employ data products of an intermediate resolution $(2 \times 1504^3$ particles) run with a box-size of $(100\,\mathrm{Mpc})^3$ referred to as REFL0100N1504 (hereafter simply EAGLE).

2.2 Galaxy Selection

In all of the aforementioned simulations, gravitationally-bound substructures are identified using a SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009) which relies on a friends-of-friends (FoF; Davis et al. 1985) algorithm to identify parent groups. In this paper, the total mass identified by SUBFIND is used for calculating global galaxy properties whereas observational works typically estimate quantities within a given aperture. We limit the sample from each simulation to central galaxies with stellar masses $8.0 < \log(M_* \ [M_\odot]) < 12.0$ and (total) gas mass[1] limits of $\log(M_{gas} \ [M_\odot]) > 8.5$. This ensures that we have $\sim 10^2$ star particles and $\sim 5 \times 10^2$ gas particles per selected galaxy, which we consider well-resolved (see Table 1 for initial baryon mass resolutions of each simulation). We note that while this resolution is too coarse to make statements about spatially resolved properties within the galaxies (e.g., metallicity gradients), it is sufficient for global properties of the systems. Further, following from Donnari et al. (2019), Pillepich et al. (2019), Nelson et al. (2021), Hemler et al. (2021), and Garcia et al. (2023), we define a specific star formation main sequence (sSFMS) in order to select only star forming galaxies. The sSFMS is defined through a linear-least squares fit to the median relation of galaxies with stellar mass $\log(M_* \ [M_\odot]) < 10.2$ in mass bins of 0.2 dex. Beyond $\log(M_* \ [M_\odot]) > 10.2$, we extrapolate the linear fit. From this sSFMS, we define (and omit) quiescent galaxies as those

---

[1] We note this gas mass threshold includes both the star forming and non-star forming gas.

whose sSFRs lie greater than 0.5 dex below this relation. We note that our key results are not sensitive to these selection criteria for the sSFMS, which we discuss further in Appendix B.

With all of these choices, at $z = 0$ we obtain 44,965 galaxies in Illustris, 22,133 in TNG, and 14,720 in EAGLE. While all of these simulations have very similar halo mass functions, we note that the overall number of galaxies in the sample is quite different between the three. The difference between the number of galaxies in EAGLE and TNG can be attributed to: (i) the larger box of TNG simply containing more galaxies and (ii) lower gas fractions in EAGLE making it more sensitive to our gas mass cut at the low mass end. Illustris, on the other hand, has an elevated stellar mass function compared to TNG, which is due to the different feedback and wind implementations (Pillepich et al. 2018a; Pillepich et al. 2018b), as such there are a factor of ∼2 more galaxies at $z = 0$ than in TNG.

Throughout all of this analysis gas-phase metallicities are from star-forming gas particles only as a fair comparison with observations, whose gas-phase metallicities are limited to emission diagnostics from star forming regions of galaxies (see Kewley & Ellison 2008; Kewley et al. 2019). Gas is determined to be star-forming based on criteria outlined in Section 2.1.1 for Illustris/TNG and Section 2.1.3 for EAGLE.

We aim to offer predictions for the high redshift ($z \lesssim 10$) MZ$_*$R in each of these simulations. However, due to the aforementioned mass limits of our sample, beyond $z = 8$ there are not enough galaxies to create a statistically significant sample. We therefore offer predictions for galaxies only out to $z = 8$.

3 RESULTS

3.1 Stellar MZR

Critically, the analysis done in this work relies on the underlying relationship between the mass of a system and the total metal content locked in stars within galaxies. To that end, we first quantify the MZ$_*$R in Illustris, TNG, and EAGLE as a baseline for the rest of the results in this work. In these three simulations we find that a relationship between stellar mass and stellar metallicity exists, as illustrated in Figure 1 for each simulation at $z = 0$. We include the MZ$_*$Rs from Gallazzi et al. (2005) and Panter et al. (2008) as a point of comparison in Figure 1; however, we caution the reader that comparing metallicities from observations and simulations is not even-handed (see Nelson et al. 2018). Specifically, computing metallicities by averaging the metal content of star particles weighted by their mass (as is done in all three simulations analysed here) produces systematically elevated stellar metallicities compared to using radiative transfer codes more akin to mock observations of galaxies (Guidi et al. 2016). In that light, it is perhaps unsurprising that the $z = 0$ MZ$_*$Rs presented in Figure 1 and higher redshift versions in Figure 2 are all elevated compared to the observations. Nevertheless, it should be noted that while Illustris and TNG appear to follow much closer to the Gallazzi et al. (2005) and Panter et al. (2008) MZ$_*$Rs compared to EAGLE, a higher resolution run of EAGLE with modified sub-grid parameters (RECAL-L025N0752) more closely reproduces the observed MZ$_*$R (see Schaye et al. 2015, their Figure 13).

In terms of the shape of the relation in the simulations, we find qualitative agreement between the three: roughly power-law with constant slope at low masses, leveling out at higher masses, qualitatively similar to the observations. In detail, however, the three simulations exhibit different behaviours. Specifically, we find that the MZ$_*$R normalisation at $z = 0$ for the lowest mass galaxies vary
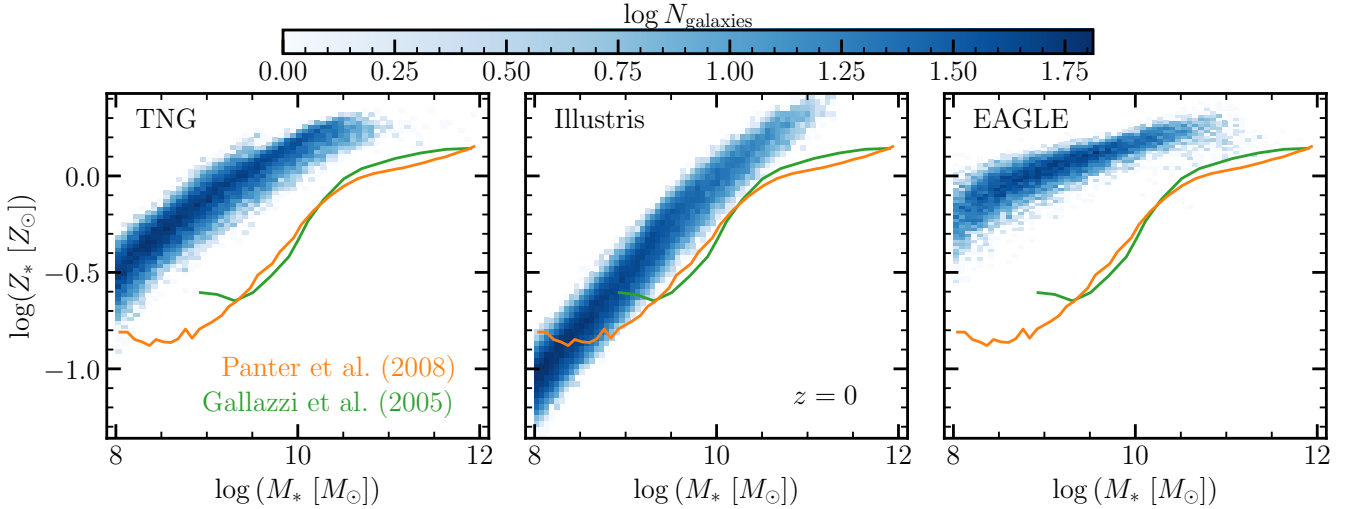
Figure 1. Stellar mass versus stellar metallicity (both scaled to solar; $Z_\odot = 0.0127$) for star forming galaxies in TNG, Illustris, and EAGLE (left, centre, and right panels, respectively) at $z = 0$. Each panel is colour-coded by the total number of galaxies within each pixel, with lighter colours corresponding to fewer galaxies. We overplot the MZ$_*$R from local observations (green, Gallazzi et al. 2005; orange, Panter et al. 2008). It is important to note that comparisons between stellar metallicities from simulations to stellar metallicities from observations is not necessarily one-to-one (see text for more details), thus we caution the reader in any comparisons against observed MZ$_*$R trends.
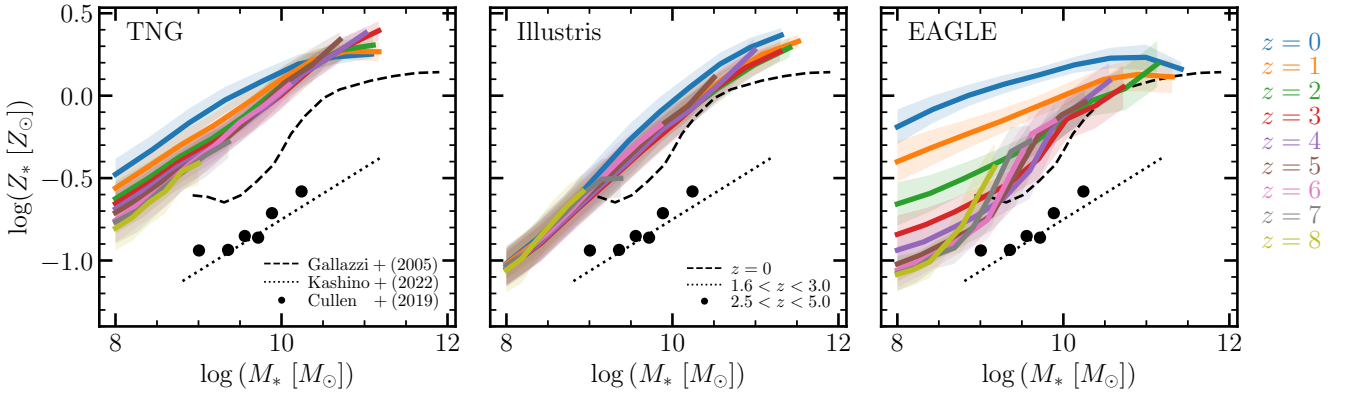
Figure 2. MZ$_*$R, as defined in Section 3.1, for star-forming galaxies in the TNG, Illustris, and EAGLE simulations (left, centre, and right panels, respectively) for $z = 0 - 8$. Galaxies are split into several different mass bins, and the median metallicity within each bin is measured. We require that there must be at least 10 galaxies in each mass bin shown (i.e., higher masses are less populated at higher redshift). The shaded regions correspond to the standard deviation about the median mass bins. The black lines and points correspond to observations from Gallazzi et al. (2005) at $z = 0$ (dashed line), Kashino et al. (2022) ranging from $1.6 < z < 3.0$ (dotted line), and Cullen et al. (2019) from $2.5 < z < 5.0$ (points).

by nearly 1.0 dex between the simulation with the EAGLE galaxies being more enriched than those in TNG, which, in turn, are more enriched than the galaxies in Illustris. At higher masses, $M_* \sim 10^{10} M_\odot$, however, the normalisation of the metallicities appears more similar between the different simulations. We also find that the slope of the power-law relation at low masses differs between the simulations: Illustris galaxies increase in metallicity more rapidly as a function of stellar mass, than those in TNG and EAGLE.

At $z > 0$, we find that differences in normalisation persist, and, in addition, the normalisations appear to evolve differently between the various simulations. In Figure 2, we trace the MZ$_*$R back to $z = 8$ for each simulation, which we compare to observations from Gallazzi et al. (2005) at $z = 0$, Kashino et al. (2022) at $1.6 < z < 3.0$, and Cullen et al. (2019) at $2.5 < z < 5.0$ (though, see above discussion

about comparison of stellar metallicities with observations). For simplicity, we define the MZ$_*$R simply as the median metallicity within fixed mass bins[2]. The shaded regions about each MZ$_*$R represents the scatter about the relation, which we define as the standard deviation of metallicities within each mass bin. Consistent with the $z = 0$ MZ$_*$R, the higher redshift Illustris galaxies are less enriched at low masses and follow a steeper trend than their TNG and EAGLE counterparts. However, around $z \gtrsim 4$, the EAGLE galaxies have changed normalisation enough to be more consistent with Illustris and there is a significant increase in the slope of the relation.

---

[2] We note that if a particular mass bin does not have more than 10 galaxies, we omit it from our sample.

To that end, there is a marked difference in the redshift evolution of the MZ_*R between the three simulations: TNG has some modest normalisation evolution with lower redshift galaxies being more enriched than high redshift, EAGLE has a similar, albeit more extreme, evolution, and Illustris shows virtually no redshift evolution. This lack of evolution in the stellar metallicities over cosmic time in Illustris has been noted previously (D'Souza & Bell 2018). These authors attribute the modest redshift evolution of the MZ_*R shown in Gallazzi et al. (2014) at $z \approx 0.7$ as evidence that, in the absence of robust predictions at higher redshifts, the lack of redshift evolution in the MZ_*R is physical. However, the predictions made here in TNG and EAGLE of an MZ_*R that evolves with redshift is in agreement with other simulation models (e.g., Ma et al. 2016; Yates et al. 2021) and more recent observations (e.g., Choi et al. 2014; Leethochawalit et al. 2018; Cullen et al. 2019; Beverage et al. 2021; Kashino et al. 2022).

3.2 Residual correlations in the scatter

Within the scatter of the MZ_gR residual correlations have been shown to exist. In particular, a secondary correlation with the star formation rate (SFR) or specific star formation rate (sSFR; SFR normalized by galaxy mass) has been seen both observationally (e.g., Ellison et al. 2008; Mannucci et al. 2010; Andrews & Martini 2013) and in simulations (e.g., De Rossi et al. 2017 in EAGLE; T18 in TNG). We find that galaxies' stellar metallicities also follow this trend: at a fixed stellar mass, galaxies with lower stellar metallicities generally have higher sSFRs (Figures 3 and 4), echoing a similar result at $z = 0$ for a higher resolution run of EAGLE by De Rossi et al. (2018; see Section 4.2 for a more in-depth comparison) . Figure 3 shows the MZ_*R for each simulation at $z = 0$ colour coded by their sSFR. Qualitatively, we find that for all simulations, save for TNG at $z = 0$, there is a clear residual correlation of the MZ_*R with sSFR. At $z = 0$ in TNG, lower mass galaxies ($M_* \lesssim 10^{9.5} M_\odot$) the residual correlation appears weaker than their Illustris or EAGLE counterparts. However, as we show in Appendix A, at $z > 0$, the galaxies fall into the more familiar relation like that of Illustris and EAGLE at $z = 0$. Given the overall model similarities – and differences – between Illustris and TNG, we attribute the lack of a clear residual correlation at $z = 0$ in TNG to the redshift-dependent winds in the model (which was not present in Illustris).

More quantitatively, Figure 4 shows the stellar metallicities as a function of sSFR in fixed mass bins (coded by their colour) for each simulation at $z = 0$. The overplotted lines are the best fit within the corresponding mass bin. We find that the majority of these best fits (at all redshifts) have negative slopes, indicating that, at a fixed stellar mass, galaxies with elevated stellar metallicities generally have lower specific star formation rates. As previously mentioned, one notable exception to this trend is the lowest mass galaxies in TNG at $z = 0$. This complements the qualitative statements above: the slope of the correlations in TNG at $z = 0$ here in the lowest mass bins ($10^{8.0} - 10^{9.0} M_\odot$) is inverted. We show in Appendix A that at higher redshift, the residual correlation is much more apparent in this mass range for TNG. Another notable exception is in the highest mass bins. In all three simulations, the highest mass bins (purple in Figure 4), appear much flatter than, or inverted with respect to, their lower-mass counterparts out to $z = 2$ for EAGLE and $z = 1$ for Illustris and TNG. Interestingly, this flattening and inversion of the residual correlation within higher stellar masses has been noted previously in both simulations (Yates et al. 2012, L-GALAXIES; De Rossi et al. 2017, EAGLE) and observations (Yates et al. 2012, Sloan Digital Sky Survey DR7) within the gas-phase metallicities and in

simulations (De Rossi et al. 2018, EAGLE) in the stellar metallicities. De Rossi et al. (2017, 2018) show, using modified AGN parameters within EAGLE, that this inversion is due to AGN feedback, which is almost certainly the case in the highest mass galaxies analysed here.

3.2.1 Projection of least scatter

Mannucci et al. (2010; henceforth M10) posit that if the metallicity, SFR, and stellar mass are all correlated, then perhaps there exists a tighter parameterisation than a traditional MZR. They fit the mass-(gas-phase) metallicity-SFR ($M_* - Z_{gas} -$ SFR) relation with a linear combination of the stellar mass and star formation rates,

$$\mu_\alpha = \log M_* - \alpha \log(\text{SFR}) , \qquad (1)$$

where α is a free parameter[3]. This free parameter encodes the relative importances of stellar mass and SFR. For example, taking $\alpha = 0.0$, the original form of the MZR is recovered, which would suggest that the SFR is unimportant in minimzing the scatter of the relation. Conversely, with $\alpha = 1.0$ the primary driving factor of the scatter is the *specific* star-formation rates (sSFR) of the galaxies. We employ this same parameterisation for the stellar mass-stellar metallicity-SFR ($M_* - Z_* -$ SFR) relation, later in this section.

To determine the best μ_α we fit a linear regression of $\log(Z_* [Z_\odot])$ versus μ_α for values of α ranging from 0 to 1. Whichever value of α corresponds to the smallest standard deviation about fixed μ_α bins from the linear fit is deemed the best fit (demonstrated in Figure 5 for EAGLE at $z = 2$). Further, we define an uncertainty of this parameter by taking any α that only varies by 1% of the minimum dispersion (see dashed gray lines in top panel of Figure 5).

Figure 6 shows the evolution of the α parameter as a function of redshift in each of the simulations with the errorbars corresponding to the aforementioned 1% uncertainty of α. We compare against observational values determined in M10 ($\alpha = 0.33$), Andrews & Martini (2013, henceforth AM13; $\alpha = 0.66$), and Curti et al. (2020, henceforth C20; $\alpha = 0.55$) for the gas-phase. Since these observational values specifically refer to the gas-phase and not stellar metallicities, we caution the reader to take comparisons against observations lightly. Another important consideration is that the relative weighting of the mass and SFR has been shown to depend strongly on the metallicity diagnostic used (AM13).

Regardless, we ultimately find a non-zero evolution in the best fit α as a function of redshift. At $z = 0$, we find relatively weak dependence on the SFR; in fact, in TNG there is virtually no dependence on the SFR. This lack of importance placed on the SFR is perhaps unsurprising in light of the lack of a significant $M_* - Z_* -$ SFR relation at that redshift in TNG (mentioned in the previous section and discussed further in Appendix A). In Illustris and TNG, α increases slightly with increasing redshift. In Eagle, on the other hand, α decreases weakly with increasing redshift. While the $M_* - Z_{gas} -$ SFR relation is oftentimes referred to as the fundamental metallicity relation (FMR; M10; AM13; C20) and has been shown to have no evolution in the relation out to $z \sim 2.5 - 3.5$ (e.g., M10; Lara-López et al. 2010), we avoid interpreting the $M_* - Z_* -$ SFR as an "FMR for stars" due to the redshift evolution seen in each of the simulations presented here. Interestingly, the redshift dependence on this relationship is also noted in Fontanot et al. (2021) with a semi-analytic model (SAM; see Section 4.2 for a more in-depth comparison

---

[3] We note that more complex relationships parameterising this residual correlation are available (e.g., Lara-López et al. 2010, 2013; Curti et al. 2020). We opt to use this specific parameterisation in view of its simplicity.
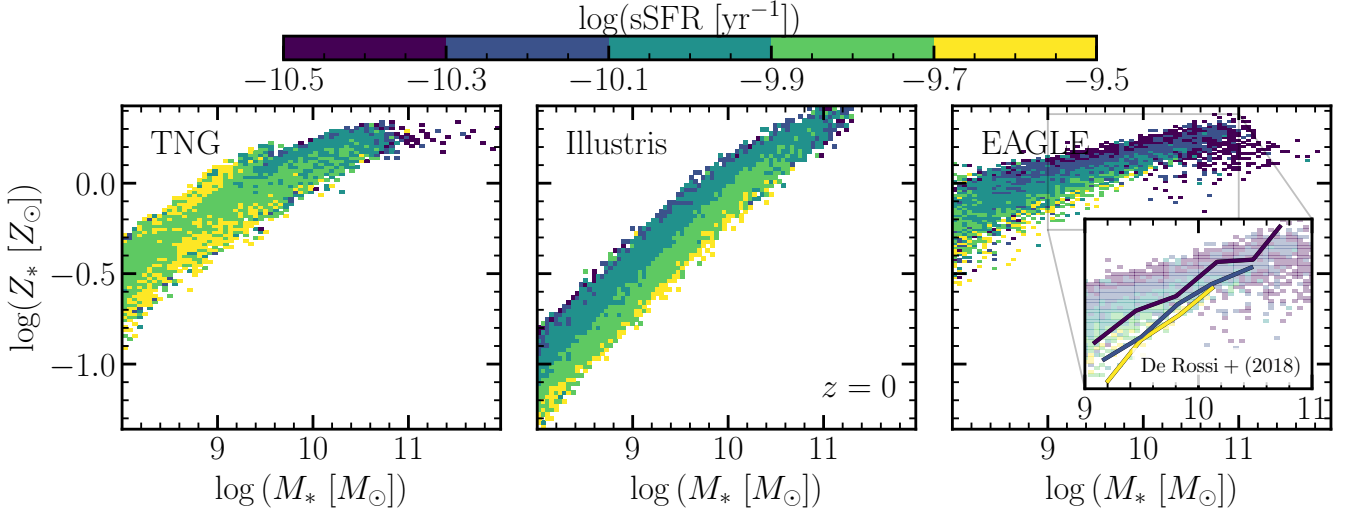
Figure 3. Same as Figure 1 colour coded by specific star formation rate (i.e., total star formation rate per unit stellar mass of the galaxy; sSFR). Lighter colours correspond to higher star forming galaxies while darker colours are lower star forming galaxies. At fixed stellar mass lower metallicity galaxies tend to have higher sSFRs. The inset on the right panel is a comparison with De Rossi et al. (2018) who use a higher resolution run of EAGLE at $z = 0$; full discussion of this comparison can be found in Section 4.2. Note that the left panel of TNG galaxies, at lower masses, does not clearly display this relationship at $z = 0$, we discuss this further in Section 3.3 and Appendix A.
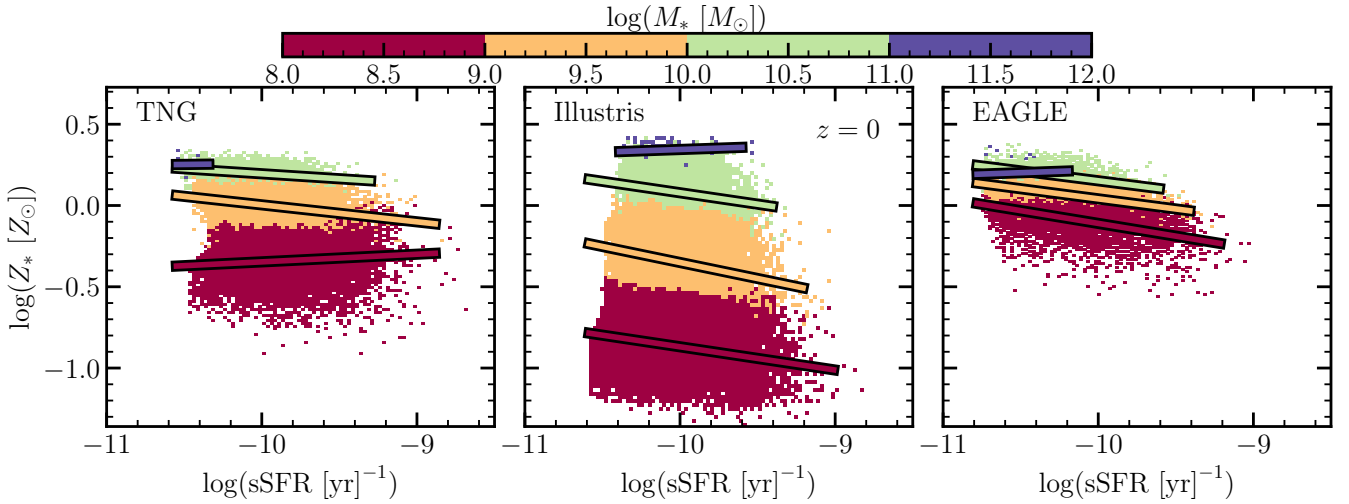
Figure 4. Stellar metallicity as a function of sSFR for star-forming galaxies in TNG, Illustris, and EAGLE (left, centre, and right, respectively) at $z = 0$ colour-coded by stellar mass bins. Galaxies are in mass bins of width 1.0 dex, centered at 8.5, 9.5, 10.5, and 11.5 $\log(M_* \, [M_\odot])$. Linear least-squares fits are provided by the outlined lines for each mass bin. We note that most of these slopes, across redshift, are negative, indicating a decrease in metallicity with increasing sSFR for a given mass.

with this model). Evidently, as we have shown, this $M_* - Z_* - \mathrm{SFR}$ relationship is not "fundamental" insofar as it evolves with time.

3.3 Origin of $M_* - Z_* - \mathrm{SFR}$ Relation

The $M_* - Z_{\mathrm{gas}} - \mathrm{SFR}$ relation (i.e., FMR) follows from fairly straightforward arguments: galactic inflows of pristine gas drive down the metallicity, but increase the available gas reservoir for star formation. Conversely, systems with limited or no gas inflows steadily consume their gas reservoirs while creating new metals. Stellar metallicities, however, are not *directly* impacted by inflows. Newly formed stars

inherit the metallicity of the gas in which they form. As such, the stellar metallicity represents the average gas-phase metallicity of the galaxy integrated over its entire lifetime. It follows, then, that the gas-phase and stellar metallicity of a system *may* be correlated, where the gas phase metallicity has the ability to more rapidly change in the presence of, e.g., rapid pristine gas inflows or significant enrichment events. In this section we investigate the extent of the correlation between gas and stellar metallicities in Illustris, TNG, and EAGLE.

By defining the MZ$_g$R in the same fashion as outlined in Section 3.1 for the MZ$_*$R and interpolating between the mass bin centers, we obtain an offset from the median relation in both the gas-phase
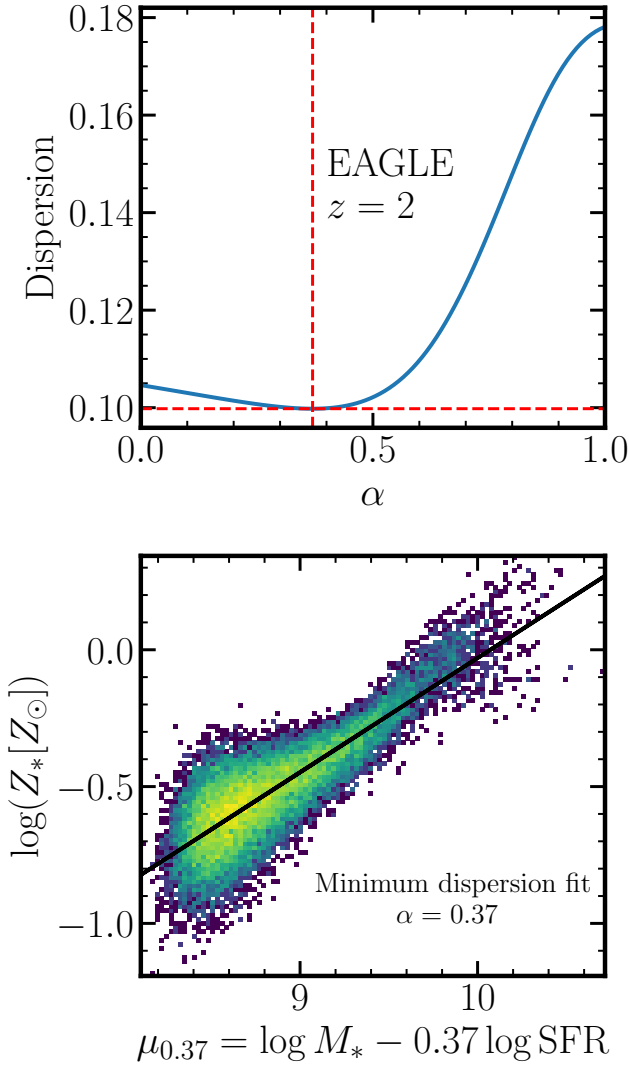
Figure 5. Demonstration of parameterisation of $M_* - Z_* -$ SFR correlation in terms of μ_α (Equation 1) following from M10. **Top:** The dispersion (i.e., standard deviation of the residuals) from the linear fit as a function of α, the free parameter in μ_α. Shown here is the dispersions for fits of the EAGLE $z = 2$ $M_* - Z_* -$ SFR relation. The red dashed lines indicate where the dispersion is minimized and at what value of α. The gray dashed lines represent the uncertainty in α values corresponding to 1% deviations from the minimum dispersion. **Bottom:** The $M_* - Z_* -$ SFR relation parameterised by the μ_α that minimizes the dispersion about the linear relation. For EAGLE at $z = 2$, we find that the best $\alpha = 0.37$.

(ΔZ_{gas}) and stellar (ΔZ_*) components for each galaxy. We find that there is a linear relationship between these offsets in all three simulations (see Figure 7 for $z = 0$ offsets). This linear relationship implies that the gas and stellar phase metallicities are tightly coupled. We fit these offsets, ΔZ_{gas} versus ΔZ_*, with a linear least squares fit requiring that the intercept pass through the origin[4]. As we discuss further in Section 4.1, slopes approaching a value of unity (like that of Illus-

---

[4] This forced intercept is sensible since the MZRs are defined as a median, thus we expect the distributions of ΔZ_* and ΔZ_{gas} to be centered around 0 (which, indeed, can be seen is roughly true from the histograms in the top panels of Figure 7).

tris' at $z = 0$) indicate a more direct relationship between the offsets of the stellar and gas phase metallicities with respect to their respective mass-metallicity relationship. Flatter slopes (e.g., EAGLE at $z = 0$) indicate a weaker relationship. Similar to the slope, understanding the strength of the correlation also depends on the magnitude of scatter within the offsets. A very tight correlation implies that stellar metallicities are more directly linked to the gas-phase, while broader distributions imply that stellar metallicities are less directly linked to the gas-phase. Thus, both the slope of the offsets and scatter within the relation inform how quickly the stars can respond to perturbations within the gas-phase metallicity. We find the slope of the offsets to be a more natural measure of the relationship here and thus use it as our primary diagnostic in this discussion.

It is important to note that while ΔZ_{gas} and ΔZ_* are related to the absolute Z_{gas} and Z_* of a system, they are different quantities. We therefore caution the reader against taking variations in the bulk offsets from the MZRs across redshifts as representative of the difference between a single galaxy's metal evolution through two cosmic times (we discuss this subtle difference with our treatment of a toy model in Section 4.1).

Performing the same offset analysis on at all redshifts, we obtain the time evolution shown in Figure 8. We find general agreement between the TNG and EAGLE simulations at $z \lesssim 3$: a shallower slope that steepens with redshift. Illustris, however, follows the opposite trend in this same redshift range, a very steep relation that weakens going back in time. Beyond $z \gtrsim 4$, all three simulations have slopes that increases with increasing redshift.

In spite of the $M_* - Z_* -$ SFR having non-negligible redshift evolution in the strength of the residual correlation across redshift, the $M_* - Z_* -$ SFR *is* rooted in the close relationship of gas-phase and stellar metallicities across time which is consistently present across time, albeit with some change in strength.

4 DISCUSSION

4.1 Toy Model

Stellar metallicities are naturally linked to gas-phase metallicities. When stars are formed, they inherit the metallicity of the gas from which they were born. Thus, changes in the gas-phase metallicity will eventually impact stellar metallicities, albeit with a time delay that accounts for the fact that it takes time for a sufficient mass of new stars to be formed with the "new" inherited gas-phase metallicity value. We model this behaviour with a simple analytic framework to better understand the origins of the MZ$_*$R and the scatter about it.

We start by defining ΔZ_* as the stellar metallicity of an individual galaxy, Z_*, subtracted from the median relation (i.e., MZ$_*$R), $\langle Z_* \rangle$, at that galaxy's particular mass. Thus, the rate of change of a galaxy's offset from the MZ$_*$R can be written as

$$\frac{\mathrm{d}(\Delta Z_*)}{\mathrm{d}t} = \frac{\mathrm{d}Z_*}{\mathrm{d}t} - \frac{\mathrm{d}\langle Z_* \rangle}{\mathrm{d}t} . \qquad (2)$$

Equation 2 tells us that the changes ΔZ_* are sensitive to not only changes in the absolute metallicity of a galaxy, but, critically, on the evolution of the normalisation of the MZ$_*$R, as well. From this we obtain two interesting regimes: (i) in the limit that the normalisation of the MZ$_*$R does not evolve, or evolves much more slowly compared to the time-scale of stellar formation (τ_{SF}), ΔZ_*'s are driven *only* by changes in the absolute metallicity and, conversely, (ii) a galaxy with no metal evolution over some time could experience a change in ΔZ_* via the overall normalisation of the MZ$_*$R changing. We begin by considering the first of the two regimes – a steady state model in
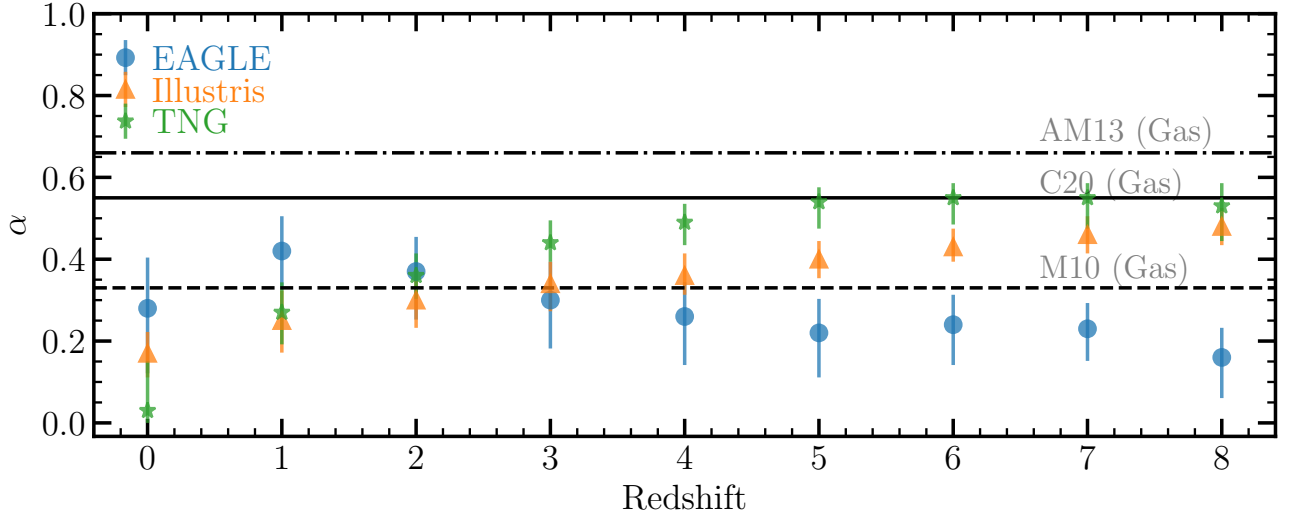
Figure 6. α, the relative weighting of star formation to stellar mass in the three component relation ($M_* - Z_* - $ SFR; as defined in Section 3.2.1), versus redshift in EAGLE (blue cirlces), Illustris (orange triangles), and TNG (green stars). The dashed line represents the relative importance found by M10 ($\alpha = 0.33$), the dot-dashed line is that found by AM13 ($\alpha = 0.66$), and the solid line is C20 ($\alpha = 0.55$; note that all these works present α with the gas-phase metals). The errorbars correspond to the uncertainty associated with 1% deviations from the minimum dispersion of the best fit μ_α (see Figure 5 and Section 3.2.1 for more details).
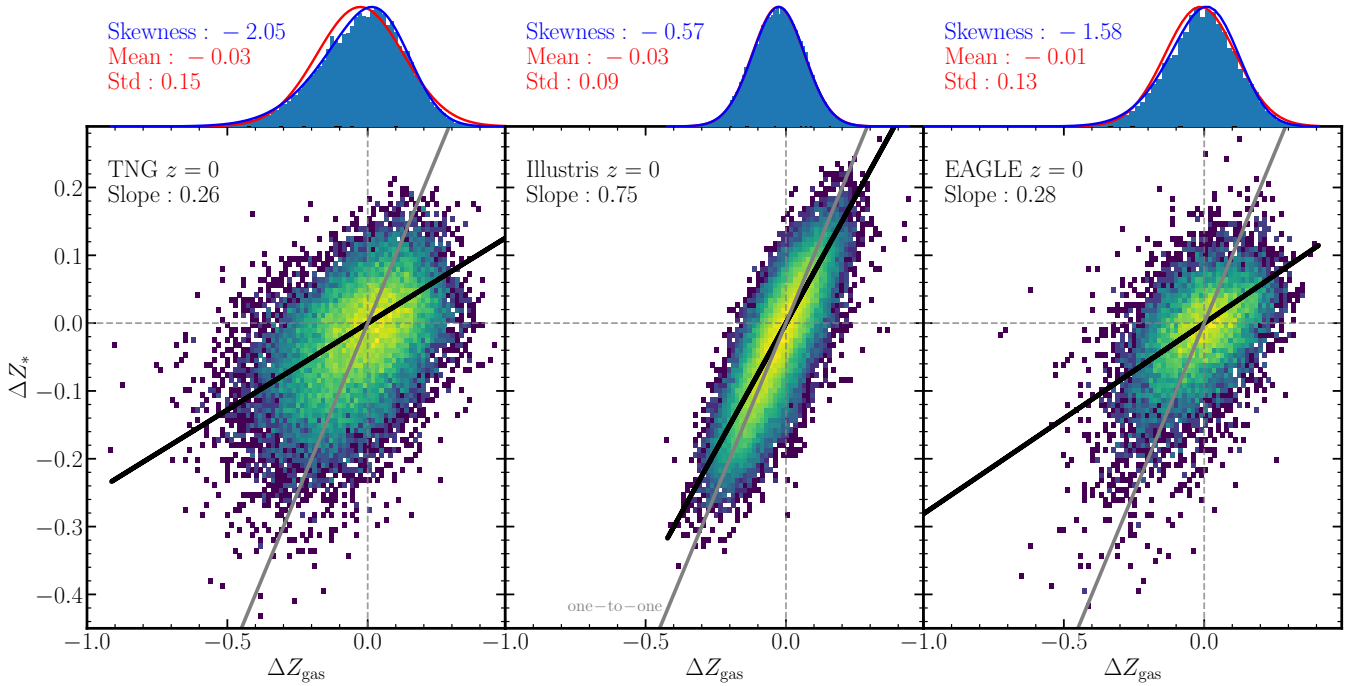
Figure 7. Bottom Panels: Offsets of both the gas-phase (abscissa) and stellar (ordinate) metallicities of star forming galaxies in EAGLE, Illustris, and TNG (left to right) at $z = 0$ from the MZRs. In each panel, a linear least squares best fit line to this relation is overplotted (black line; fixing the intercept to the origin) as well as a one-to-one line (gray solid line). The slope of the linear least squares fit is the relevant quantity in this relation and allows us to quantify how correlated the two metallicities are to each other; a slope of unity would suggest that the gas-phase and stellar metals are exactly correlated, while a slope of 0 suggests that they are uncorrelated. **Top Panels:** Normalised distributions of the offsets from each simulation's MZ$_g$R. We fit both a Gaussian distribution (red) and skewed Gaussian (blue) to the distribution. The typical mean and standard deviation of these distributions (across redshift) is ~0.0 dex and ~0.12 dex, respectively. We find that the distributions have some non-zero skew, which we discuss in more detail in Section 4.1.
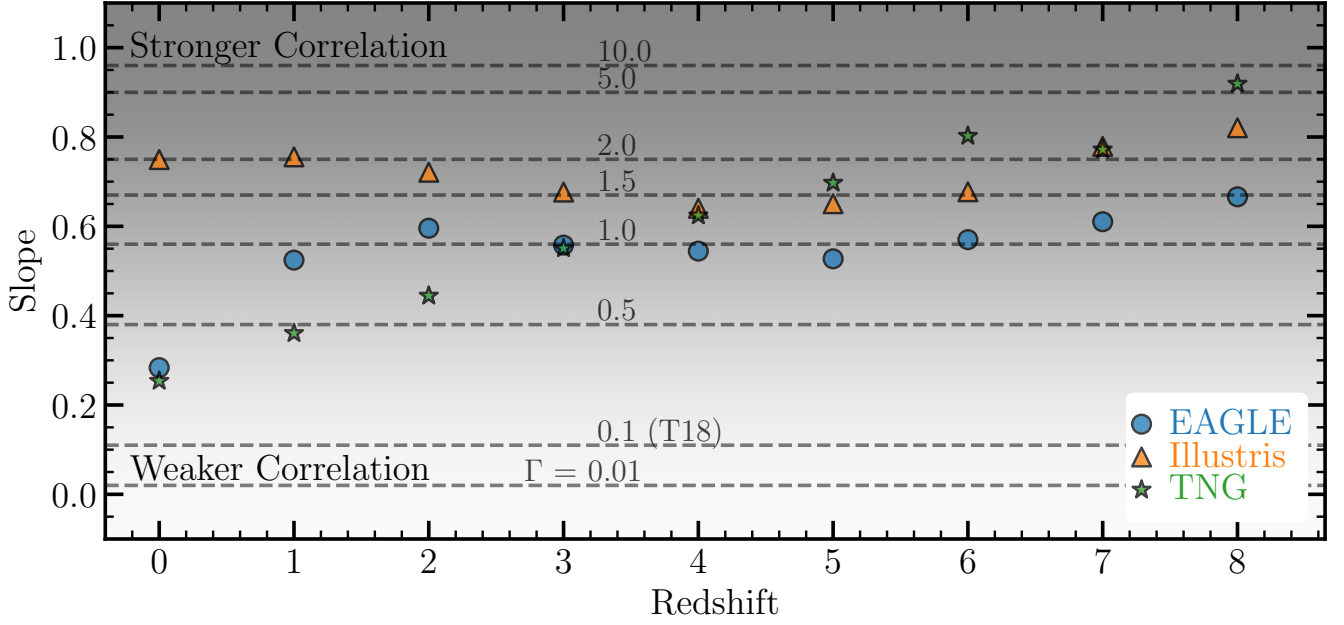
Figure 8. Slope of the offsets from the MZRs for EAGLE (blue cirlces), Illustris (orange triangles), and TNG (green stars) as a function of redshift with horizontal lines corresponding slopes obtained by running our analytic model (Section 4.1) with certain values of Γ (equation 7). The gradient background indicates that slopes closer to unity indicate a stronger relationship between $\Delta Z_{\rm gas}$ and ΔZ_*. We note that the $\Gamma = 0.1$ line corresponds to predictions by T18.
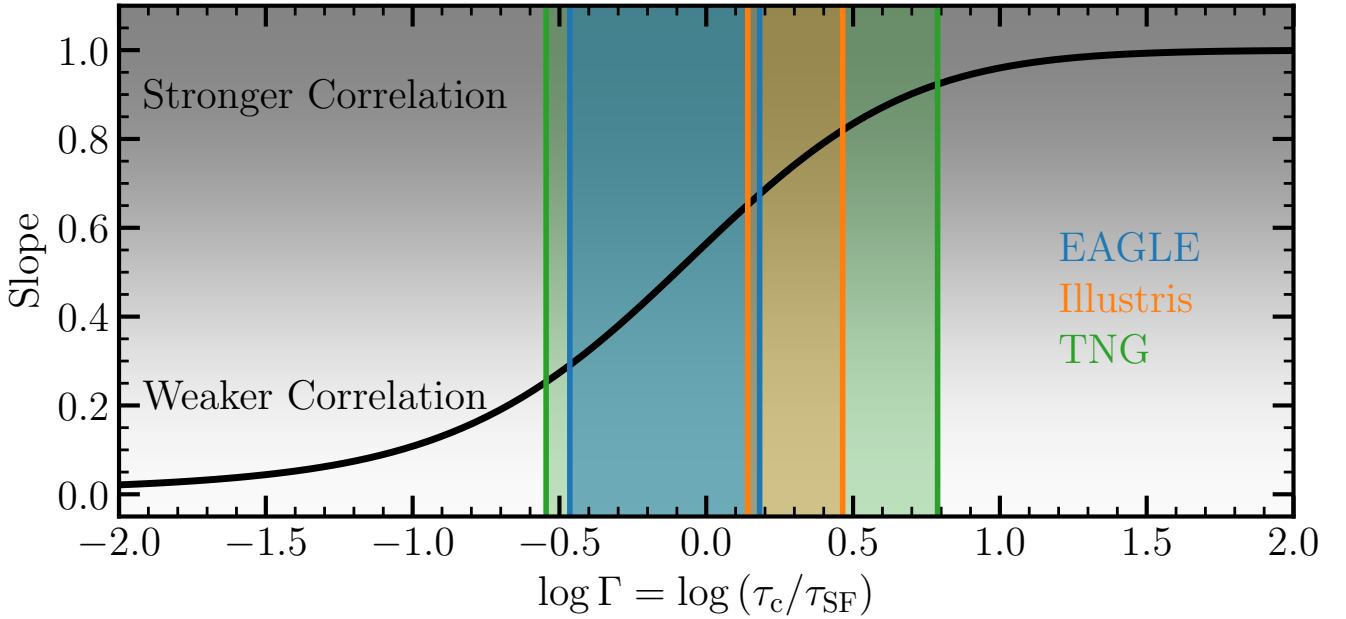
Figure 9. Slopes of the offset relation of the gas-phase and stellar metallicities (as in Figure 7) versus $\log \Gamma$, the ratio of the coherence time-scale ($\tau_{\rm c}$; the time-scale on which gas-phase metallicities are perturbed) and stellar formation time-scale ($\tau_{\rm SF}$; the inverse sSFR). The black line represents predictions by our toy model (Section 4.1). Slopes closer to unity represent a tighter correlation between the gas-phase and stellar metallicities. The shaded regions correspond to the ranges of expected values of $\log \Gamma$ for EAGLE (blue), Illustris (orange), and TNG (green) based on the slopes in Figure 8.

which there is no MZ$_*$R evolution – to better understand how a toy galaxy's gas-phase and stellar metallicities interact. Then we explore the second regime and discuss the implications of the co-evolution of the MZ$_*$R as a whole underneath galaxy-by-galaxy evolution.

4.1.1 Steady State Model

In principle, the rate of change of the stellar metallicity of an individual system should be a function of (i) the amount of stars being formed, (ii) the present gas-phase metallicity of gas forming new stars, and (iii) the present stellar metallicity. To attain an analytic relation that is characterised by all three, we first define the stellar metallicity of the galaxy as the total mass of metals locked within stars, normalized by the total stellar mass:

$$Z_* = \frac{M_{Z,*}}{M_*} \; . \tag{3}$$

By differentiating equation 3, and understanding that the rate of change of metals in stars is simply the gas-phase metallicity scaled by the amount (by mass) of stars that are being formed[5], we obtain a straight-forward expression for the rate of change of the stellar metallicity that depends on exactly the three parameters described above,

$$\frac{dZ_*}{dt} = \frac{1}{M_*}\frac{dM_*}{dt}\left(Z_{\mathrm{gas}} - Z_*\right) \; . \tag{4}$$

From equation 4, it follows that the stellar metallicity should tend towards the gas-phase metallicity value (of the star forming gas) over time. Moreover, the rate at which this occurs is set only by the amount of new stars being formed normalized by the total stellar mass of the system, i.e. the specific star formation rate.

However, the quantity that we are comparing to in this model is the *offsets* in stellar metallicity, ΔZ_*. We must therefore use equation 2 to model how the offsets in stellar metallicity respond. Substituting the absolute metallicity with its offset and MZR terms, as well as rearranging, we are left with

$$\begin{aligned}\frac{d(\Delta Z_*)}{dt} =& \frac{1}{M_*}\frac{dM_*}{dt}\left[\langle Z_{\mathrm{gas}}\rangle - \langle Z_*\rangle\right] \\ &+ \frac{1}{M_*}\frac{dM_*}{dt}\left(\Delta Z_{\mathrm{gas}} - \Delta Z_*\right) - \frac{d\langle Z_*\rangle}{dt} \; .\end{aligned} \tag{5}$$

For now we will assume that the MZ$_*$R does not evolve ($d\langle Z_*\rangle/dt = 0$) and that the stellar and gas-phase MZRs are at the same value ($\langle Z_{\mathrm{gas}}\rangle = \langle Z_*\rangle$) such that the first and third term of the left-hand side are zero. We will revisit these assumptions in Section 4.1.2. Applying these assumptions to equation 5 we are left with

$$\frac{d(\Delta Z_*)}{dt} = \frac{1}{M_*}\frac{dM_*}{dt}\left(\Delta Z_{\mathrm{gas}} - \Delta Z_*\right) \; . \tag{6}$$

With this representation of the evolution of galaxy metal content, we model the response of a toy galaxy in this paradigm. For simplicity, we assume that our toy galaxy initially lies directly on both the MZ$_g$R and MZ$_*$R, so that its offsets from either are zero[6]. Next, we draw a deviation for the gas-phase metallicity (ΔZ_{gas}) from a normal distribution with a standard deviation of 0.12 dex, corresponding to

a typical standard deviation of the distribution of gas-phase metallicity offsets in the simulations (see, e.g., top panels of Figure 7), centered at an offset of 0 dex. The toy galaxy's gas-phase metallicity tends towards the randomly drawn value over a time-scale drawn from an exponential distribution with some coherence time-scale, τ_{c}, nominally set to one-tenth a Hubble time (as per T18; see discussion later in Section 4.1.3). Physically these perturbations are meant to mimic a potential pathway that a galaxy's gas-phase metallicity will undertake as it evolves. As its gas-phase metallicity changes, we model the response of the toy galaxy's stellar metallicities by equation 6. Specifically, we model its formation of new stars with a stellar formation time-scale, τ_{SF}, as the inverse of sSFRs consistent with observations (one over the hubble time).

Simply put, the coherence time-scale represents how quickly the gas-phase metallicity is perturbed, while the stellar formation time-scale informs how quickly the stars can respond to these changes. Together the coherence and star formation time-scales fully parameterise the behavior of our toy model. Thus, it is informative to define a (dimensionless) ratio of the two:

$$\Gamma = \frac{\tau_{\mathrm{c}}}{\tau_{\mathrm{SF}}} \; . \tag{7}$$

Since, nominally, τ_{c} is one-tenth a Hubble time and τ_{SF} is a Hubble time, Γ takes a fiducial value of 0.1. It is worth noting that this Γ holds all of the predictive power of our model. By increasing Γ, the gas-phase metallicity perturbations take longer (or identically, the system forms stars more quickly) and thus the stellar metallicity might track the gas-phase metallicity closely. Conversely, by decreasing Γ, we expect that the stellar metallicities take longer to equilibrate to the gas-phase metallicities, and therefore the two metallicities are not likely to remain matched as the gas phase metallicity changes. Thus, the strength of the relation of the offsets of the metallicities from their respective MZR's is a direct effect of Γ.

By scaling the initial conditions for the sSFR[7] we find the strength (i.e., slope) of the offsets from the MZR values can be uniquely described by a value of Γ (see Figure 9). Lower (higher) Γ corresponds to a weaker (stronger) correlation between the two metallicities. This result agrees with naïve expectations: if the gas-phase metallicity evolves very quickly compared to the time it takes to form new stars, then the stellar metallicities will significantly lag behind the gas-phase metallicities.

We do note that the distribution of gas-phase metallicity offsets in the simulations more closely follows a skew-normal distribution with skewness parameters ranging from -1 to -3.5 (with the exceptions of TNG $z = 2$, Illustris $z = 0$, and EAGLE $z = 7$ and 8, which have skews of 0.0, -0.59, +0.36, and +0.70, respectively). Despite the magnitude of skew > 1 indicating that the distribution is non-Gaussian, when sampling gas-phase metallicity perturbations from a skew-normal distribution in our analytic model, we find that the overall slope of the correlation remains unchanged. Further, as mentioned in Section 3.3, we forced the intercept of the offset relation to the origin; however, when sampling from a skewed normal distribution, we remove this requirement. This is sensible since the skewed normal distribution potentially shifts the 'typical' value of the distribution away from an offset of 0. We note, however, that even if we continue to require the regression to pass through the origin, the quantitative change in the slope is modest. Thus, since the key results found here are not

---

[5] We note that to be complete there should be some exchange of metals injected back into the ISM from the stars from supernovae and AGB winds (see, e.g., Peng & Maiolino 2014). For simplicity, and to isolate its effect, we neglect this term here. We discuss the implications of this in Section 4.1.3.
[6] We note that non-zero offsets in either the gas-phase or stellar metallicities (or both) do not significantly impact the results of this toy model.

[7] Though we choose to vary sSFR, this is, in principle, exactly equivalent to decreasing the coherence time-scale. We choose the former for the sake of convenience.

significantly impacted by using a skew normal distribution, we opt to use a regular Gaussian distribution.

With the predictions offered by our analytic model and the slopes found in Section 3.3, we predict the range of values for Γ in each simulation (Figure 9). For EAGLE, we find that the $\log \Gamma$ values range from ~ -0.5 to 0.2, in TNG we find $\log \Gamma$ ranges from ~ -0.55 to 0.8, and in Illustris they range from ~ 0.2 to 0.5.

4.1.2 Impact of the underlying MZRs

It is important to note that the slopes quoted in Figures 7 and 8 are based on ΔZ_{gas} and ΔZ_* – offsets from median relations. As we noted at the beginning of this discussion, changes in ΔZ are sensitive to not only changes in a system's absolute Z but to changes in the underlying MZRs as well. Yet, in the previous section, we assume that the median relation does not evolve with time and that the gas-phase median relation is equal to that of the stellar relation. We relax that assumption in this section and consider how the evolution and interplay of the underlying relations impact the interpretation of our toy model.

Both the first and third terms on the right-hand side of equation 5 imply that a galaxy's ΔZ_* changes owing to bulk changes in the underlying median relation. While these terms are related, each represents slightly different physical mechanisms. The first term expresses the difference between the median MZ_gR and median MZ_*R. This term is qualitatively similar to that of the offset term presented in Section 4.1.1: the entire MZ_*R tends towards the MZ_gR over some time-scale. The larger the difference in the two median relations, the faster the MZ_*R will respond. Physically (both globally and individually), this links to the idea that the stellar metallicity is adopted from the gas where stars form. The next generation of stars will have a higher metallicity when the gas-phase metals are higher than that of the stars. On the other hand, the third term models the evolution of the normalisation of the MZ_*R. Individual ΔZ_*s will experience changes owing to the evolution in the MZ_*R, regardless of what the absolute stellar metallicity is doing. Changes in the MZ_*R are set partially by the whole relation tending towards the MZ_gR, but also include newly formed metals within the stars of the galaxies themselves. Importantly, both the first and third terms are sensitive to changes in the MZRs.

The evolution of MZ_*R plays a more explicit role in interpreting our toy model, we therefore isolate its effect here by only considering the second and third terms on the right-hand side of equation 5. The evolution of the normalisation only becomes important when it is significantly faster than stellar formation time-scales, however. Once MZ_*R evolution becomes significant enough, we can no longer neglect these additional terms in equation 5. By taking the difference in metallicity within all the mass bins within the MZ_*R at each redshift and dividing by difference in Hubble times between two given redshits, we find that the median evolution of the normalisation of the MZ_*R is 0.08 dex/Gyr in EAGLE, 0.02 dex/Gyr in TNG, and 0.0 dex/Gyr in Illustris. We note that these values are a crude approximation that do not account for possible mass and redshift dependencies on the evolution rate of the MZ_*R; nevertheless, they are useful for considering the typical redshift evolution within each simulation. Furthermore, these evolution rates reiterate the qualitative interpretations of Figure 2 in Section 3.1 – the most redshift evolution of the MZ_*R is found within EAGLE, with modest evolution in TNG, and virtually no evolution in Illustris.

When adding the MZ_*R evolutionary term with both the Illustris and TNG rates, we find no significant effect on the overall obtained slope of the relation of ΔZ_{gas} versus ΔZ_* compared to the slope

from the steady state model. For EAGLE, since the MZ_*R normalisation changes more quickly, the evolutionary term becomes more important. Specifically, when using the EAGLE MZ_*R evolution rate, $\log \Gamma \lesssim -0.25$ (corresponding to only $z = 0$) is where the model begins to deviate significantly from the steady state. In this regime, the MZ_*R is evolving so rapidly compared to the stellar formation time-scales that the second term in Equation 2 dominates the change in ΔZ_* over time. The significant MZ_*R evolution has the net effect of systematically moving the average ΔZ_* away from 0. Yet, by definition, these ΔZ_* values should be centered around ~ 0. The reason behind the systematic shift is that our model only runs a single toy galaxy's metallicity track; the additional evolutionary term just represents the *typical* amount the MZ_*R evolves – the model does not actually recalculate an MZ_*R based on a population of galaxies' stellar metallicities at each time step. As such, fitting the relation with a line that passes through the origin (as we did in Section 3.3) produces ill-defined results. In order to rectify the toy model in this regime, a more sophisticated approach would be required that traces evolutionary tracks of multiple systems at once all while evolving the overall normalisation of the MZ_*R by a fixed amount at each time step, from which ΔZ_*'s from the median relation could be calculated. While it is possible to construct such a model, we favor the simplicity of the toy model detailed above in light of the relatively good job that it does reproducing the simulations' offsets in conjunction with the limited range in which such a careful tracing of MZ_*R evolution is necessary.

The first term on the right hand-side of equation 5 explicitly encodes information about the difference in absolute stellar and gas-phase MZRs but also implicitly encodes information about their respective evolution. In the limit that the neither MZR evolves, this term is just a constant that changes the rate at which ΔZ_* evolves. A constant, non-zero offset should steepen the slope of ΔZ_{gas} versus ΔZ_* for $\langle Z_{gas} \rangle > \langle Z_* \rangle$ as it increases the rate of change of ΔZ_*. It should be noted that for large differences in $\langle Z_{gas} \rangle$ and $\langle Z_* \rangle$ this term may play a significant role. Both the MZ_*R and MZ_gR do have some evolution associated with them, however. The median gas-phase metallicity evolutionary rates are 0.09 dex/Gyr in EAGLE, 0.08 dex/Gyr in TNG, and 0.05 dex/Gyr in Illustris (all faster than their MZ_*R counterparts). A complete treatment of this model would need to explicitly track the absolute metallicities of the MZRs, as well as their evolution during the toy galaxy's evolution. Since the MZ_gR evolutionary values are all significant, the current model presented here does not have the capabilities to perform such a task (for reasons detailed above).

Besides the impact on the change on $\langle Z_{gas} \rangle$, MZ_gR evolution has no impact on our toy model. The reason that evolution of the MZ_gR is otherwise unimportant in our model is that we draw the gas-phase metallicity perturbations about the MZ_gR itself (ΔZ_{gas} values), centered at an offset of 0. Moreover, we find that the scatter about the MZ_gR does not vary significantly across time in the simulations. Therefore, our perturbations to the gas-phase metallicity are not sensitive to changes of the MZ_gR.

Finally, we note that the underlying MZRs have some structure as a function of mass. We make the simplifying assumption that the stellar mass of the galaxy does not evolve significantly with time. In reality, there would likely be some contribution to ΔZ_* owing to the galaxy increasing in stellar mass as it assembles its stellar populations.

In summary, the additional MZ_*R evolutionary term may appear necessary to properly compare our toy model's predictions to the obtained slopes of the offsets from Section 3.3; however, with the exception of $z = 0$ in EAGLE, there is no difference in the interpreta-

tion by including the secondary term. Regardless, we caution that the significant MZR evolution in EAGLE makes the previously derived values of $\log \Gamma$ more uncertain compared to, e.g., Illustris which has very little MZR evolution. Additionally, we find that the difference in the median MZRs increases the slope of ΔZ_{gas} versus ΔZ_* when $\langle Z_{\text{gas}} \rangle > \langle Z_* \rangle$ and is a constant offset.

4.1.3 Limitations of the model

We note that there appears to be some tension with the T18 model, which (effectively) predicts that $\log \Gamma = -1$ for TNG. Recall that we neglect the return of mass and metals back into the ISM from stars dying. This removes a $(1 - R)$ term on the sSFR, where R is the mass return fraction from stars (identically, a $1/(1 - R)$ scaling on τ_{SF}). If we instead included this return fraction, assuming that is a constant over τ_c, it is equivalent to just scaling Γ by $(1 - R)$. This does indeed shift us closer to agreement with T18, the extent of the agreement, however, depends on the adopted value of R. In order to produce a slope that corresponds to a $\log \Gamma = -1$, we must assume stellar mass return values ranging from $\sim 66 - 97\%$. This value is much larger than current understanding predicts (i.e., $\lesssim 40\%$; Pillepich et al. 2018a; Hopkins et al. 2023). It is also possible that T18 underestimates the amount of time that a galaxy will persist above or below the MZR. The equivalent coherence time-scale from T18 (called metallicity evolution time-scale in that work) uses a Pearson correlation coefficient to describe the strength of the correlations in the offsets from the MZgR. After one coherence (metallicity evolution) time-scale the strength of the correlation is weaker, but not gone entirely. Indeed, van Loon et al. (2021) find that the time-scale for evolution of gas-phase metallicities is $\gg 1$ Gyr in EAGLE, which (particularly at high redshift) is significantly greater than a tenth of the Hubble time. Thus, the subtle difference in interpretation of this time-scale may push the coherence time-scale (as defined in this work) to larger values, shifting the T18 predictions closer towards agreement with this work. Furthermore, the stellar formation time-scale could also be dependent on the mass of the system, as Matthee & Schaye (2019) show, with less massive galaxies having more short time-scale fluctuations. Another possibility is that our method of randomizing metallicities isn't really a fair match to what's happening in the simulations. For example, we model the gas-phase metallicity changing at a constant rate over the entire coherence time-scale. The rate of change of gas-phase metallicity might change as a function of time (from, e.g., infalling material with a non-uniform density distribution). Moreover, it is possible that after a single coherence time-scale the system's gas-phase metallicity is not *immediately* perturbed again, giving additional time for the stellar metallicity value to tend towards the gas-phase value.

It is also worth noting that this model does not include every physical mechanism with which a system's metal content could evolve. For one, the role that mergers play is underestimated by our treatment here. While the role that infalling gas is accounted for (i.e., a possible mechanism for the gas coherence time-scale), merging galaxies also bring in stars that could change the stellar metallicity completely disjoint from new star formation. Dry mergers in particular would bypass both the gas-phase coherence and stellar formation time-scales entirely by changing the stellar metallicity of a system very quickly without forming new stars or perturbing the gas-phase metallicity. The fraction of stars that are born ex-situ (outside of the galaxy) depends on the mass of the galaxy (Rodriguez-Gomez et al. 2016; Davison et al. 2020). The ex-situ fraction is negligible for a majority of the sample analysed in this work (see, e.g., Pillepich et al. 2015) and therefore its impact on the overall stellar metallicity is likely negligible. Past $\log M_* \gtrsim 10.2 \log M_\odot$ this fraction becomes more significant, however ($> 25\%$). Additionally, AGN feedback plays a vital role in the evolution of metal content of high mass galaxies, yet does not fit into our toy model here. De Rossi et al. (2017, 2018) and van Loon et al. (2021), using the EAGLE model, show that AGN have the effect of suppressing both the gas-phase and stellar metal content in the highest mass galaxies via quenching star formation and ejecting enriched material with outflows. These effects are thought to cause the inversion of the $M_* - Z_{\text{gas}} - \text{SFR}$ relation at the highest masses (as in Yates et al. 2012), and are likely contributing to the inversion of the $M_* - Z_* - \text{SFR}$ relation in the highest mass galaxies shown in Figure 4. Regardless, such quenching or outflow mechanisms are simply not taken into account in this toy model, yet it is clear that they are important to some extent in the metal evolution of the highest mass systems.

4.2 Comparisons with other models

Just as in observations, there is limited work quantifying the $M_* - Z_* - \text{SFR}$ relation in simulations. Fontanot et al. (2021) used the semi-analytic model GAlaxy Evolution and Assembly (GAEA) and found a residual trend between $M_* - Z_* - \text{SFR}$. These authors investigate the evolution of both the gas-phase and stellar metallicities through time and recover both the typical gas-phase FMR and a similar relation for stars. Similar to the results shown in this work (see Section 3.2; Figure 6), they find the $M_* - Z_* - \text{SFR}$ correlation evolves with redshift. Differently than this work, however, they quantify this change in the relation not in terms of a three component relation, but instead by measuring an offset from the $z = 0$ defined $M_* - Z_* - \text{SFR}$ relation [8]. As such, this comparison is limited to systems with overlapping masses or sSFRs at $z = 0$ and the desired redshift ($z = 2.24, 4.18,$ and 4.88). In these systems, they find that offsets from the $z = 0$ $M_* - Z_* - \text{SFR}$ relation are a factor of $\gtrsim 2$ more than in the gas-phase. These authors consider this increased scatter evidence for significant redshift evolution of the $M_* - Z_* - \text{SFR}$ and attribute this evolution to the increase of baryonic mass locked in the stars as galaxies evolve.

De Rossi et al. (2018) also investigated the correlation of gas content within the MZ$_*$R within EAGLE (though a smaller box, higher-resolution run; RECAL-L025N0752, $m_{\text{baryon}} = 0.226 \times 10^6 M_\odot$) at $z = 0$. Those authors found that the scatter of the MZ$_*$R at $z = 0$ is correlated with not just the sSFR, but also the gas fraction and mass-weighted age of a galaxy's stellar population, specifically that higher metallicity systems: (i) are older, (ii) have a lower gas fractions, and (iii) have lower sSFRs (and vice-versa). The inset in the right-most panel of Figure 3 shows the quantitative comparison between the high resolution EAGLE and our analysis here. Qualitatively, these three findings corroborate what we find in Section 3.2 with the sSFR of galaxies. While the EAGLE run we analyse has slightly elevated metalicities compared to the higher resolution run of De Rossi et al. (2018), the overall relationship is remarkably similar between the two runs, suggesting the $M_* - Z_* - \text{SFR}$ relationship is not dependent on the resolution of the simulation and rather a feature of the model at least at this low redshift. However, more robust comparisons at varying resolutions and at higher redshifts are required to confirm this.

---

[8] Fontanot et al. (2021) refer to this as the "FMR" in their work. We refrain from using this language here to be self-consistent with previous sections of this work (see Figure 6 and Section 3.2.1 for more details).

4.2.1 Implications for stellar feedback modeling

As mentioned in Section 2.1, all of the cosmological simulations presented in this work are of the sub-grid ISM pressurization type (i.e., SH03 for Illustris and TNG, and SDV08 for EAGLE) which are characterized by gradual (i.e. relatively non-bursty) star formation and stellar feedback. Thus, it appears that the existence of a $M_* - Z_* - $ SFR correlation in all of these models – which is naturally explained by our toy model with smoothly evolving star formation rates and regularly evolving metallicities – constitutes a reasonably generic prediction of these smooth (i.e. non-bursty) stellar feedback models. These sub-grid implementations of the star forming ISM are not the only ways to model the ISM, however. The Feedback in Realistic Environments (FIRE; Hopkins et al. 2014) model, for one, attempts to more explicitly model the multi-phase ISM. As a consequence, the star formation (and, as a result, stellar feedback) in the FIRE model can occur in large bursts – especially in low-mass and high-redshift systems. These large bursts of feedback can quickly remove a significant fraction of the gas from the galaxies.

The analytic model presented in Section 4.1 is implicitly smooth star formation: gas is accreted onto the galaxy and, over some stellar formation time-scale τ_{SF}, forms into stars. While this model is certainly simplistic, it produces a sensible result in the stellar metallicity of the system lags behind the gas-phase metallicity as subsequent generations of stars form from the enriched ISM, evolve, and return the metals to form new stars. Our model breaks down in a bursty star formation paradigm. Rapid ejections of gas from the discs reduces the time that stars have to form as the star forming gas is removed from the system. It is likely that these bursts would significantly disrupt the process of stellar metallicites catching up to the gas-phase values. Thus, the strength of the residual correlation within the MZ$_*$R, even if such a thing exists, is likely to be sensitive to the "burstiness" of stellar feedback within galaxies. Further, these interruptions would likely leave their mark on the correlation between the gas-phase and stellar metallicities, as well, potentially weakening the overall correlation.

Interestingly, the feedback prescription used in the SAM employed in Fontanot et al. (2021; GAEA, Hirschmann et al. 2016) takes some inspiration from the FIRE model (Hopkins et al. 2014) insofar as it attempts to model bursty feedback, yet, as previously mentioned, Fontanot et al. (2021) still recover a similar residual correlation within the MZ$_*$R. While quantitative comparisons on the strength of the correlation are difficult to make, owing to the largely different methodology in the two works, it is worth noting that the relationship indeed does exist to *some* extent in more bursty feedback models. However, SAMs reconstruct feedback via approximations to observations combined with theoretical models. Thus, it is possible that these approximations might curtail the effectiveness of the feedback compared to hydrodynamic simulations and a residual correlation may or may not appear in full hydrodynamic simulations, e.g., FIRE.

Further examination, in these hydrodynamic simulations with bursty feedback is required to fully understand the effect of burstiness on this residual trend. However, we suggest that the (lack of) correlation between gas and stellar metallicities for galaxy populations may provide a promising opportunity for constraining or confirming the role of bursty feedback in galaxy assembly.

5 CONCLUSIONS

In this work, we select star-forming, central galaxies with stellar masses $8.0 < \log(M_* \ [M_\odot]) < 12.0$ and gas masses $\log(M_{gas} \ [M_\odot]) > 8.5$ from the Illustris, TNG, and EAGLE cosmological simulations. Taking the global metallicity of these galaxies, we create a stellar mass-stellar metallicity relation (MZ$_*$R) and stellar mass-gas-phase metallicity relation (MZ$_g$R) for each at $z = 0 - 8$. With these, we can find each galaxy's offset from each MZR and examine residual trends. We examine these residual trends by using a simple analytic model to track the co-evolution of the gas-phase and stellar metallicities.

Our conclusions are as follows:

• The general trend of each of the MZ$_*$R in each simulation is similar: less metals in low mass galaxies which increases with stellar mass before plateauing at the highest masses. The overall normalisation of the MZ$_*$R (Figure 1) and the redshift evolution (Figure 2), however, vary significantly in between Illustris, TNG, and EAGLE.

• The scatter about the MZ$_*$R is correlated with the sSFR of the galaxies. Higher (lower) metallicity galaxies tend to have lower (higher) specific star formation rates (Figure 3 and 4). This result follows suit from observations of the scatter about the MZ$_g$R (e.g., Ellison et al. 2008; M10; Lara-López et al. 2010). However, we find that the dependence on SFR varies as a function of redshift (Figure 6) in contrast to observations and results from simulations of an "FMR" in the gas-phase.

• The origin of both this correlated scatter and the MZ$_*$R can be traced back to the intercorrelation of stellar and gas-phase metallicities, which we have fairly strong correlations within these simulations (Figures 7 and 8).

• We present a simple toy model for the causal evolution of the stellar metallicities from the gas-phase metallicities. We determine that the driving factor between the correlation of the gas-phase and stellar metallicities depends on the interplay of the coherence time-scale (how quickly gas-phase metallicities change) and the stellar formation time-scale (how quickly new stars form). By definition a dimensionless parameter, Γ, as the ratio of these two time-scales, we can quantify the correlation between the gas-phase and stellar metallicities in the simulations we analyse (Figure 9).

The broad-strokes agreement in the residual correlations between Illustris, TNG, and EAGLE suggest that this $M_* - Z_* - $ SFR may be a fairly generic prediction of galaxy formation models with this ISM pressure support creating smooth star formation histories. As the gas-phase metallicities are perturbed, the stellar metallicities have sufficient time to respond and "catch-up" to the gas-phase metals. This, however, is not obviously true *a priori* in models with more bursty assembly histories, thus the existence of this residual correlation could potentially provide an observable diagnostic for ISM feedback models.

DATA AVAILABILITY

All reduced data products, analysis scripts, and figures the support the conclusions of this work are all available publicly at https://github.com/AlexGarcia623/interplay_gas_stars. Data from the Illustris and IllustrisTNG is already available on each project's respective website. Illustris: https://www.illustris-project.org/data/ and IllustrisTNG: https://www.tng-project.org/data/. Similarly, data products from the EAGLE simulations are available for public download via the Virgo consortium's website: https://icc.dur.ac.uk/Eagle/database.php

REFERENCES

Andrews B. H., Martini P., 2013, ApJ, 765, 140
Baker W. M., et al., 2023, MNRAS, 519, 1149
Beverage A. G., Kriek M., Conroy C., Bezanson R., Franx M., van der Wel A., 2021, ApJ, 917, L1
Blanc G. A., Lu Y., Benson A., Katsianis A., Barraza M., 2019, ApJ, 877, 6
Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Cicone C., 2013, MNRAS, 433, 1425
Bothwell M. S., Maiolino R., Peng Y., Cicone C., Griffith H., Wagg J., 2016, MNRAS, 455, 1156
Chabrier G., 2003, PASP, 115, 763
Chisholm J., Rigby J. R., Bayliss M., Berg D. A., Dahle H., Gladders M., Sharon K., 2019, ApJ, 882, 182
Choi J., Conroy C., Moustakas J., Graves G. J., Holden B. P., Brodwin M., Brown M. J. I., van Dokkum P. G., 2014, ApJ, 792, 95
Crain R. A., et al., 2015, MNRAS, 450, 1937
Cresci G., Mannucci F., Curti M., 2019, A&A, 627, A42
Cullen F., et al., 2019, MNRAS, 487, 2038
Cullen F., et al., 2021, MNRAS, 505, 903
Curti M., Mannucci F., Cresci G., Maiolino R., 2020, MNRAS, 491, 944
Curti M., et al., 2023, arXiv e-prints, p. arXiv:2304.08516
D'Souza R., Bell E. F., 2018, MNRAS, 474, 5300
Dalcanton J. J., 2007, ApJ, 658, 941
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Davison T. A., Norris M. A., Pfeffer J. L., Davies J. J., Crain R. A., 2020, MNRAS, 497, 81
De Rossi M. E., Bower R. G., Font A. S., Schaye J., Theuns T., 2017, MNRAS, 472, 3354
De Rossi M. E., Bower R. G., Font A. S., Schaye T., 2018, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 60, 121
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Donnari M., et al., 2019, MNRAS, 485, 4817
Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, ApJ, 672, L107
Elmegreen B. G., 1999, ApJ, 527, 266
Faucher-Giguère C.-A., 2018, MNRAS, 473, 3717
Fontanot F., et al., 2021, MNRAS, 504, 4481
Fraser-McKelvie A., et al., 2022, MNRAS, 510, 320
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Gallazzi A., Bell E. F., Zibetti S., Brinchmann J., Kelson D. D., 2014, ApJ, 788, 72
Garcia A. M., et al., 2023, MNRAS, 519, 4716
Genel S., et al., 2014, MNRAS, 445, 175
González Delgado R. M., et al., 2014, ApJ, 791, L16
Greener M. J., et al., 2022, MNRAS, 516, 1275
Guidi G., Scannapieco C., Walcher J., Gallazzi A., 2016, MNRAS, 462, 2046
Halliday C., et al., 2008, A&A, 479, 417

Hemler Z. S., et al., 2021, MNRAS, 506, 3024
Hirschmann M., De Lucia G., Fontanot F., 2016, MNRAS, 461, 1760
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Hopkins P. F., et al., 2023, MNRAS, 519, 3154
Huang C., et al., 2019, ApJ, 886, 31
Kashino D., et al., 2022, ApJ, 925, 82
Kennicutt Robert C. J., 1998, ApJ, 498, 541
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 425, 2027
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511
Kobayashi C., Karakas A. I., Lugaro M., 2020, ApJ, 900, 179
Kriek M., et al., 2019, ApJ, 880, L31
Langeroodi D., et al., 2022, arXiv e-prints, p. arXiv:2212.02491
Lara-López M. A., et al., 2010, A&A, 521, L53
Lara-López M. A., López-Sánchez Á. R., Hopkins A. M., 2013, ApJ, 764, 178
Lazar A., et al., 2020, MNRAS, 497, 2393
Lee H., Skillman E. D., Cannon J. M., Jackson D. C., Gehrz R. D., Polomski E. F., Woodward C. E., 2006, ApJ, 647, 970
Leethochawalit N., Kirby E. N., Moran S. M., Ellis R. S., Treu T., 2018, ApJ, 856, 15
Lian J., Thomas D., Maraston C., Goddard D., Comparat J., Gonzalez-Perez V., Ventura P., 2018, MNRAS, 474, 1143
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
Maiolino R., et al., 2008, A&A, 488, 463
Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115
Marinacci F., et al., 2018, MNRAS, 480, 5113
Matthee J., Schaye J., 2019, MNRAS, 484, 915
McAlpine S., et al., 2016, Astronomy and Computing, 15, 72
Naiman J. P., et al., 2018, MNRAS, 477, 1206
Nelson D., et al., 2018, MNRAS, 475, 624
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
Panter B., Jimenez R., Heavens A. F., Charlot S., 2008, MNRAS, 391, 1117
Pasquali A., Gallazzi A., van den Bosch F. C., 2012, MNRAS, 425, 273
Patrício V., et al., 2016, MNRAS, 456, 4191
Peng Y.-j., Maiolino R., 2014, MNRAS, 443, 3643
Pillepich A., Madau P., Mayer L., 2015, ApJ, 799, 184
Pillepich A., et al., 2018a, MNRAS, 473, 4077
Pillepich A., et al., 2018b, MNRAS, 475, 648
Pillepich A., et al., 2019, MNRAS, 490, 3196
Rodriguez-Gomez V., et al., 2016, MNRAS, 458, 2371
Sanders R. L., Shapley A. E., Topping M. W., Reddy N. A., Brammer G. B., 2023, arXiv e-prints, p. arXiv:2303.08149
Savaglio S., et al., 2005, ApJ, 635, 260
Schaller M., Dalla Vecchia C., Schaye J., Bower R. G., Theuns T., Crain R. A., Furlong M., McCarthy I. G., 2015, MNRAS, 454, 2277
Schaye J., 2004, ApJ, 609, 667
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J., et al., 2015, MNRAS, 446, 521
Schmidt M., 1959, ApJ, 129, 243
Sextl E., Kudritzki R.-P., Zahid H. J., Ho I.-T., 2023, arXiv e-prints, p. arXiv:2303.11024
Shapley A. E., Reddy N. A., Sanders R. L., Topping M. W., Brammer G. B., 2023, arXiv e-prints, p. arXiv:2303.00410
Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999
Sommariva V., Mannucci F., Cresci G., Maiolino R., Marconi A., Nagao T., Baroni A., Grazian A., 2012, A&A, 539, A136
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., White M., Hernquist L., 2001, ApJ, 549, 681

Springel V., et al., 2018, MNRAS, 475, 676

Topping M. W., Shapley A. E., Reddy N. A., Sanders R. L., Coil A. L., Kriek M., Mobasher B., Siana B., 2020, MNRAS, 499, 1652

Torrey P., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 427, 2224

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985

Torrey P., et al., 2018, MNRAS, 477, L16

Torrey P., et al., 2019, MNRAS, 484, 5587

Tremonti C. A., et al., 2004, ApJ, 613, 898

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769

Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, A&ARv, 28, 2

Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 425, 3024

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., et al., 2014b, Nature, 509, 177

Weinberger R., et al., 2017, MNRAS, 465, 3291

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574

Yang N., Scholte D., Saintonge A., 2022, arXiv e-prints, p. arXiv:2212.10657

Yates R. M., Kauffmann G., Guo Q., 2012, MNRAS, 422, 215

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

Zahid H. J., Kewley L. J., Bresolin F., 2011, ApJ, 730, 137

Zahid H. J., Kudritzki R.-P., Conroy C., Andrews B., Ho I. T., 2017, ApJ, 847, 18

van Loon M. L., Mitchell P. D., Schaye J., 2021, MNRAS, 504, 4817

APPENDIX A: REDSHIFT EVOLUTION IN TNG

As mentioned in Section 3.2.1 (and shown in Figures 3 and 6), at redshift 0 in TNG, there is virtually no tertiary dependence on SFR within the MZ$_*$R (i.e., $\alpha \sim 0.0$). Whereas in Figure 3 there is no clear trend of decreasing SFR with increasing metallicity in galaxies of stellar mass $\log(M_* \, [M_\odot]) \lesssim 9.5$ in TNG, the left-most panel of Figure A1 shows a much clearer trend with galaxies' SFRs. Furthermore, the left panel of Figure A2 shows that we recover (mostly[9]) negative slopes within the relationship between the sSFR and stellar metallicity at fixed stellar mass.

One possible source of the lack of a clear relationship in TNG at $z = 0$ is the different implementations of galactic scale winds in the updated TNG models (see Pillepich et al. 2018a, for a complete description of the differences between the models). The original Illustris framework launches the winds as velocities dependent on the local dark matter velocity dispersion. TNG takes this one step further and introduces a redshift dependence to the winds by setting a velocity floor for wind injection. This velocity floor ensures that low mass haloes do not have unphysical mass loading factors. Consequently, the low redshift star formation suppression and high redshift feedback are more effective in TNG compared to Illustirs. These different wind implementations give rise to different mass-metallicity relations that evolve differently with time. It is possible that this increased star formation suppression at low redshifts specifically plays a part in the lack of an $M_* - Z_* - $ SFR relation at $z = 0$ in TNG.

This idea is further corroborated by investigating TNG50, a (51.7 Mpc)^3 box run of TNG (Pillepich et al. 2019). Following all of the same selection criteria from Section 2.2, we perform a similar

analysis in TNG50-1 and TNG50-2, the highest resolution runs of TNG50. TNG50-1 has an initial baryon mass resolution of $8.4 \times 10^4 \, M_\odot$ which is much higher resolution than TNG100-1. TNG50-2's initial baryon mass resolution is more comparable to that of TNG100-1 at $6.7 \times 10^5 \, M_\odot$. We find a similar lack of dependence on sSFR at $z = 0$ in both TNG50-1 and 50-2 (top panels of Figure A3). The sSFR dependence, however, appears at $z = 1$ in TNG50-1 and 50-2 (bottom panels of Figure A3). This suggests that the MZ$_*$R's lack of a dependence on sSFR at $z = 0$ is likely a feature of the model and not a by-product of the mass resolution of TNG100-1.

It is interesting to note that a similar lack of correlation can be seen in Torrey et al. (2019; top left panel of their Figure 7) in the gas-phase metallicity.

## APPENDIX B: DEPENDENCE ON THE SPECIFIC STAR FORMATION MAIN SEQUENCE SELECTION CRITERIA

When selecting galaxies for our sample, we restrict ourselves to only star forming galaxies characterised by a sSFMS (see Section 2.2). One of the key characteristics of the selection is that any galaxy that is significantly below the relation is classified as quiescent and omitted from the sample. The fiducial classification is that any galaxy that is 0.5 dex below the sSFMS is omitted, in this appendix we consider three alterations: (i) any galaxy greater than 0.1 dex below the sSFMS is omitted, (ii) any galaxy greater than 1.0 dex below the sSFMS is omitted, and (iii) only galaxies not forming stars at all are omitted.

Using the each definition of the sSFMS, the left panels of Figure B1 shows the determination of α (as in Figure 6, Section 3.2.1) as a function of redshift in EAGLE, TNG, and Illustris in the top, middle, and bottom panels, respectively. We find that across redshift α is relatively insensitive to changes in the sample, further the *quantitative* trends described in the main text are unchanged. Similarly, the right panels of Figure B1 shows the slopes of the offset relations between $\Delta Z_{\rm gas}$ and ΔZ_* (as in Figure 8, Section 3.3) as a function of redshift for the the four different sample criteria in EAGLE, TNG, and Illustris in the top, middle, and bottom panels, respectively. In this we also find that variations to the sample criteria produce only modest changes in the quantitative values of the slopes and the qualitative trends are largely preserved. Thus, we conclude that our key results are fairly insensitive to the sSFMS criteria outlined in Section 2.2.
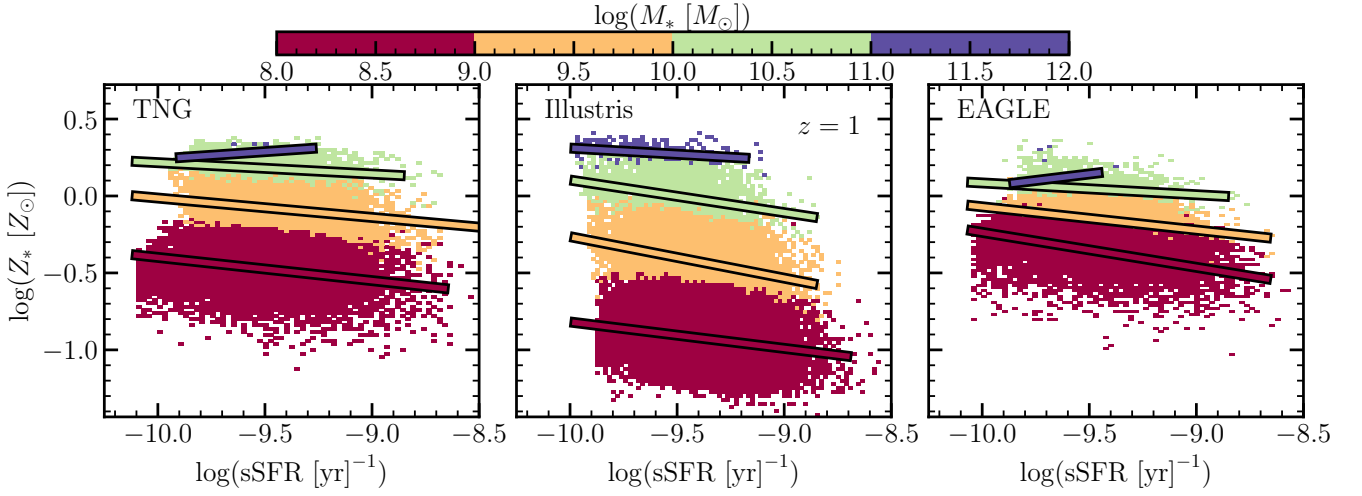
This paper has been typeset from a TeX/LaTeX file prepared by the author.

---

[9] As mentioned in Section 3.2, the highest mass bins have flatter (or even inverted) relationships; similar to the findings of Yates et al. (2012).
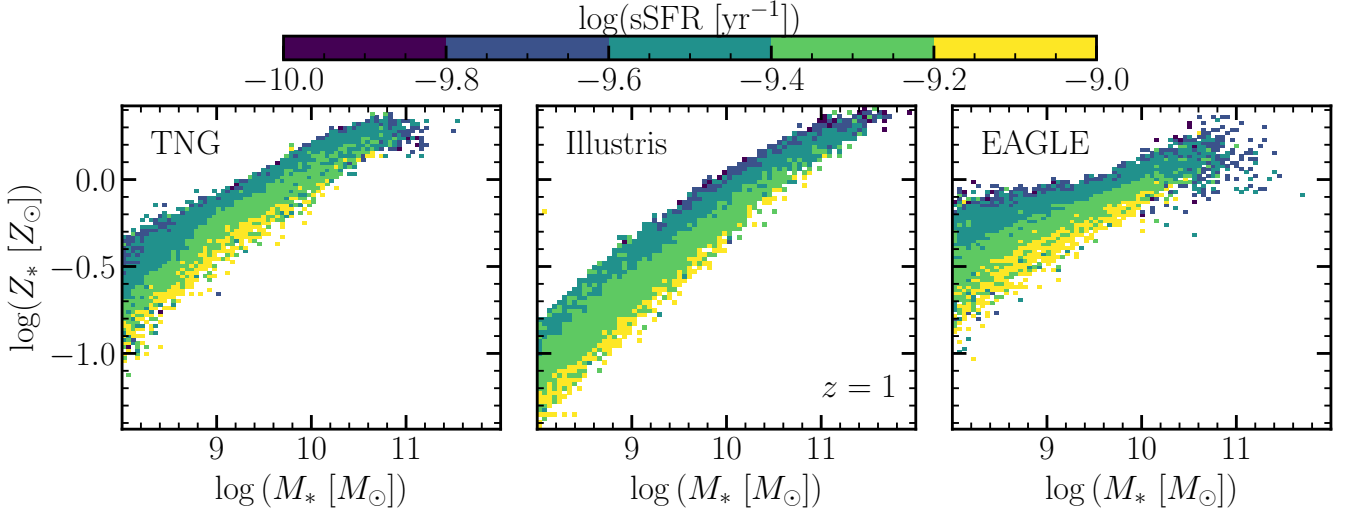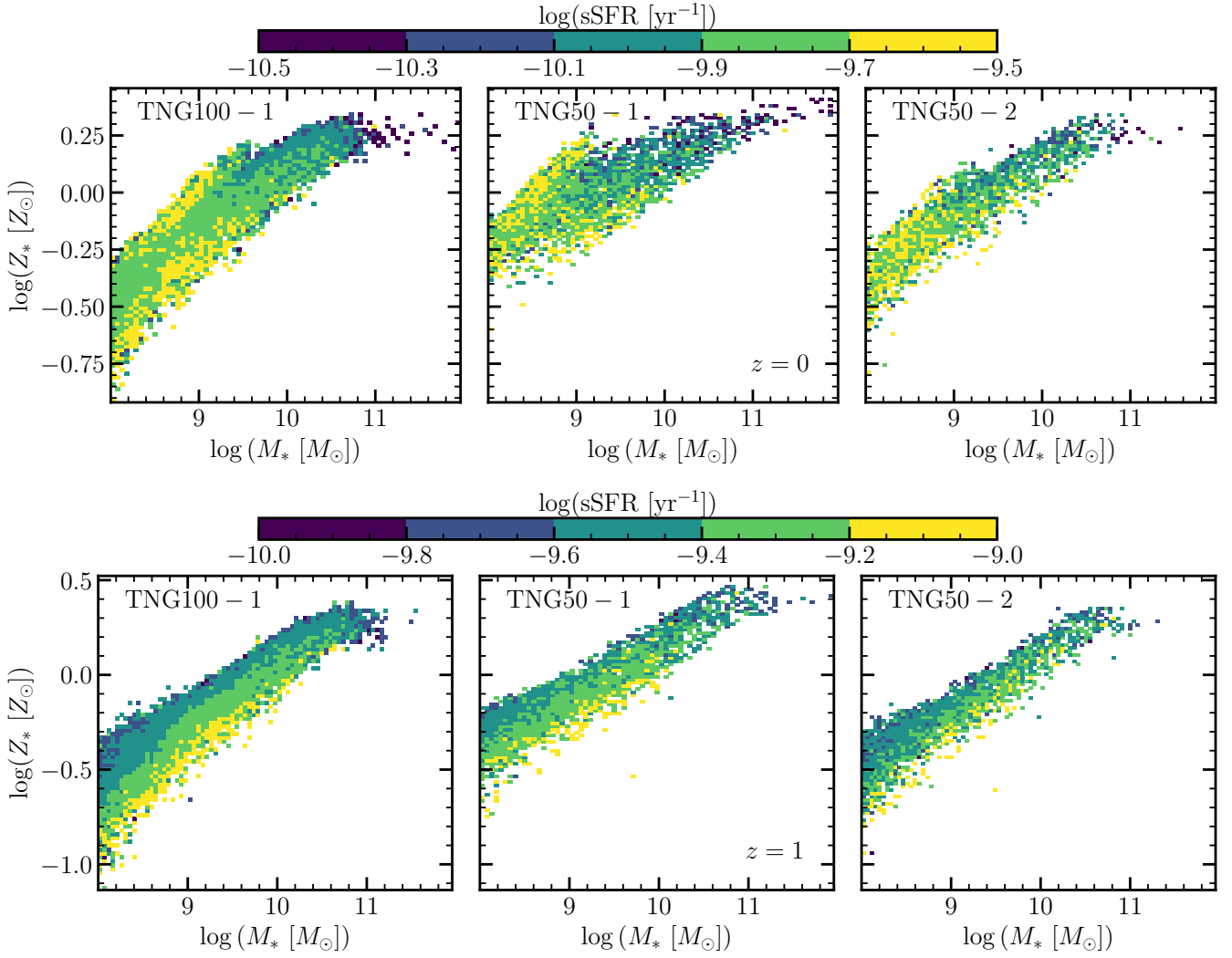
Figure A1. Same as Figure 3, but at $z = 1$. Note the clear evidence of a $M_* - Z_* -$ SFR relation at lower masses in the left panel for TNG.

Figure A2. Same as Figure 4, but at $z = 1$. Note the clear evidence of a $M_* - Z_* -$ SFR relation at lower masses in the left panel for TNG.

Figure A3. Same as Figure 3 for (left to right) TNG100-1 (called TNG in previous plots), TNG50-1, and TNG50-2. The top panels are at $z = 0$ and the bottom panels are at $z = 1$.
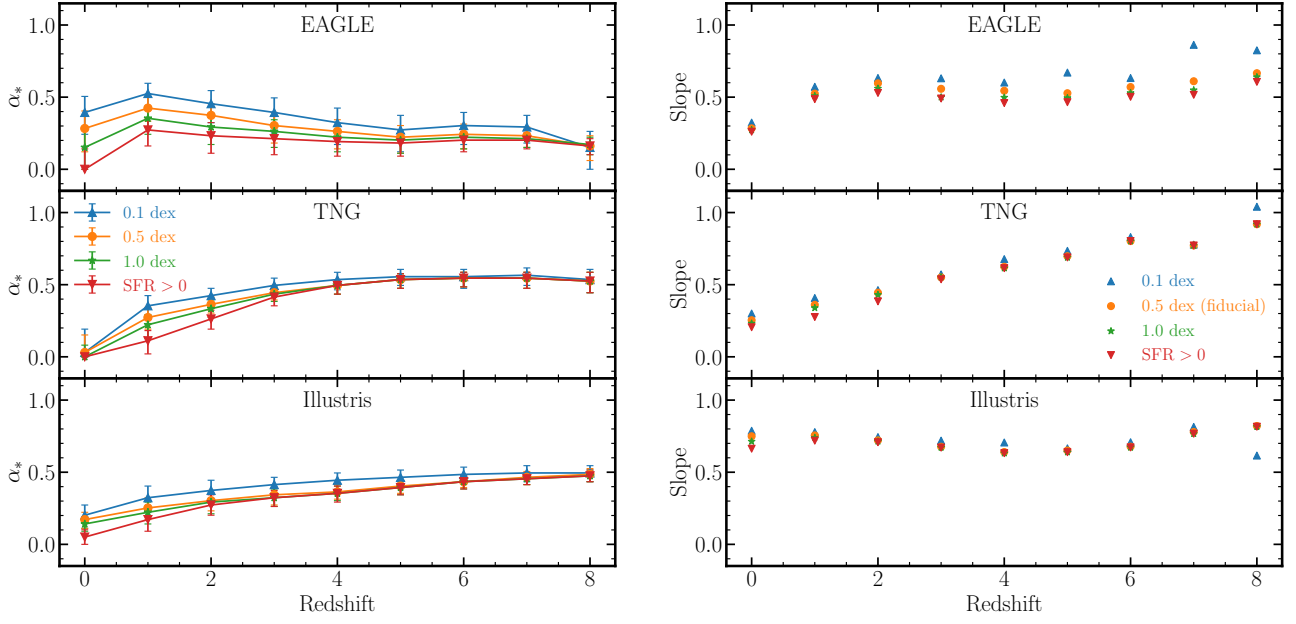
Figure B1. Left: Determination of α, the free parameter in Equation 1, for EAGLE, TNG, and Illustris (top, middle, bottom, respectively) as a function of redshift for the four sSFMS selection variations. **Right:** Value for the slope of the offset relations (as in Figure 7) for EAGLE, TNG, and Illustris (top, middle, bottom, respectively) as a function of redshift for the four sSFMS selection variations.