The stellar chemical abundances of simulated massive galaxies at $z = 2$

Jee-Ho Kim [ORCID],[1]* Sirio Belli [ORCID][2] and Rainer Weinberger [ORCID][3]

[1]*Department of Astronomy, University of Virginia, 530 McCormick Road, Charlottesville, VA 22904-4325, USA*
[2]*Dipartimento di Fisica e Astronomia, Università di Bologna, via Gobetti 93/2, I-40122 Bologna, Italy*
[3]*Canadian Institute for Theoretical Astrophysics, 60 St George Street, Toronto, ON M5S 3H8, Canada*

ABSTRACT

We analyse the stellar abundances of massive galaxies ($\log M_*/M_\odot > 10.5$) at redshift, $z = 2$, in the IllustrisTNG simulation with the goal of guiding the interpretation of current and future observations, particularly from *JWST*. We find that the effective size, R_e, of galaxies strongly affects the abundance measurements: both [Mg/H] and [Fe/H] are anticorrelated with R_e, while the relative abundance [Mg/Fe] slightly increases with R_e. The α enhancement as tracked by [Mg/Fe] traces the formation time-scale of a galaxy weakly, and mostly depends on R_e. Aperture effects are important: measuring the stellar abundances within 1 kpc instead of within R_e can make a large difference. These results are all due to a nearly universal, steeply declining stellar abundance profile that does not scale with galaxy size – Small galaxies appear metal-rich because their stars live in the inner part of the profile where abundances are high. The slope of this profile is mostly set by the gas-phase abundance profile and not substantially modified by stellar age gradients. The gas-phase abundance profile, in turn, is determined by the strong radial dependence of the gas fraction and star-formation efficiency. We develop a simple model to describe the chemical enrichment, in which each radial bin of a galaxy is treated as an independent closed-box system. This model reproduces the gas-phase abundance profile of simulated galaxies, but not the detailed distribution of their stellar abundances, for which gas and/or metal transport are likely needed.

Key words: software: simulations – galaxies: evolution – galaxies: formation – galaxies: stellar content.

1 INTRODUCTION

One of the most important properties of galaxies is their chemical composition, which can be measured independently for gas and stars. While the *gas-phase metallicity* is sensitive to processes such as star formation, inflows, and outflows (e.g. Maiolino & Mannucci 2019), the *stellar metallicity* probes the past history of a galaxy, since stars inherit the metallicity of the gas from which they form. For example, the stellar metallicity of quiescent galaxies can be used to investigate their quenching, a process that happened several Gyrs earlier (Peng, Maiolino & Cochrane 2015). Thus, observational and theoretical studies of stellar metallicity can provide crucial insights on the formation and evolution of galaxies.

Observations in the local Universe have shown that more massive galaxies tend to have a higher metal content; this is true whether one considers gas-phase metallicity (Tremonti et al. 2004) or stellar metallicity (Gallazzi et al. 2005). Multiple theories have been developed to explain this mass–metallicity relation (see Maiolino & Mannucci 2019, and references therein). In a popular scenario, first proposed by Larson (1974), massive galaxies are more metal-rich because their deep potential wells are able to retain most of the metals ejected by supernova-driven winds, while in low-mass systems these winds are more effective at removing metals. Other scenarios

include additional important ingredients such as gas inflows and star-formation efficiency (e.g. Dalcanton 2007).

More recently, it has become possible to study chemical compositions using cosmological simulations of galaxy formation (Finlator & Davé 2008; De Rossi et al. 2015; Bahé et al. 2017; Trayford & Schaye 2019; Hemler et al. 2021), which treat processes self-consistently – e.g. star formation and cosmological environment activity including galaxy mergers, gas inflows, chemical enrichment, gas flows induced by stellar feedback, and transport of different chemical elements. These simulations have been used to understand the emergence of gas-phase scaling relations (Torrey et al. 2019; Henry et al. 2021; van Loon, Mitchell & Schaye 2021) and the properties of local galaxies (Christensen et al. 2018; Grand et al. 2018, 2019), with some studies also discussing the stellar metal abundances in galaxy populations (Guidi et al. 2016; Ma et al. 2016; Taylor & Kobayashi 2016).

In addition to measuring the overall metal content, analysing the abundance of individual elements can provide further constraints on the formation and evolution of galaxies. In particular, α elements such as oxygen and magnesium are mostly produced by core-collapse supernovae, which occur soon after the formation of short-lived massive stars ($M_* > 8\,M_\odot$); while iron peak elements like iron and nickel are produced by type Ia supernovae, which occur up to several Gyr after the onset of star formation, as they form from lower mass stars with longer life spans. Thus, galaxies that formed rapidly must feature large amounts of α elements in their stars compared to iron-peak elements. Measurements of the relative abundance [α/Fe] = $\log (N_\alpha/N_{Fe}) - \log (N_\alpha/N_{Fe})_\odot$ in the local Universe revealed that

* E-mail: jk3sxp@virginia.edu

massive galaxies are more α-enhanced, suggesting that they formed more rapidly than lower mass galaxies (Trager et al. 2000; Thomas et al. 2010). Theoretical studies indicate a fundamental plane of stellar mass, star-formation rate, and α enhancement in star-forming galaxies (Matthee & Schaye 2018).

Studies of local galaxies are ultimately limited by the difficulty of distinguishing between the stars that formed *in situ* and those that were accreted from satellite galaxies. In order to disentangle this effect, it is necessary to directly observe the progenitors of massive galaxies at high redshift. However, obtaining the high-quality spectra that are required to measure stellar abundances is particularly challenging for distant galaxies. Measurements of α enhancement for representative samples have only recently become available at intermediate redshift ($z \sim 0.7$, Beverage et al. 2021; Borghi et al. 2022), while at $z \sim 1$–2, only a few individual measurements exist (Kriek et al. 2016, 2019; Jafariyazani et al. 2020; Lonoce et al. 2020). This observational landscape is expected to change substantially thanks to the spectroscopic capabilities of *JWST*, which will open up a new era for the study of stellar abundances and their evolution.

The interpretation of metal abundance observations is further complicated by the presence of radial gradients in galaxies. In the local Universe, metallicity declines with Galactocentric distance (e.g. González Delgado et al. 2015), while [α/Fe] may feature a weak increase with radius (e.g. van Dokkum et al. 2017). The meaning of galaxy-integrated abundance measurements, therefore, can be different for galaxies with different gradients, sizes, or morphologies (Zibetti & Gallazzi 2022).

The main goal of this work is to use numerical simulations to guide the interpretation of stellar metallicity observations, particularly for galaxies at the peak of cosmic star-formation history, at $z \sim 2$. We aim to connect *spatially resolved* abundance gradients, which are closely related to the physical processes shaping the formation and evolution of galaxies, to *galaxy-wide* measurements, which are easier to obtain for distant galaxies. We focus on massive systems, which are key to understanding the assembly and formation of massive haloes in the early Universe.

The paper is laid out as follows: Section 2 describes the IllustrisTNG (TNG) numerical simulation and the data used in the analysis; Section 3 explores galaxy-wide stellar abundances, while Section 4 focuses on the spatially resolved abundance distributions and the connection between stellar and gas metallicity. We discuss our results in Section 5 and give a summary in Section 6.

2 DATA

2.1 Simulation

We analyse predictions of stellar metallicity from the TNG project (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018b; Springel et al. 2018). The TNG project is the next-generation simulation suite after the Illustris simulation (Genel et al. 2014; Vogelsberger et al. 2014a, b), using updated physics and modelling methods to expand our understanding of galaxy formation processes. TNG incorporates gravity, ideal magnetohydrodynamics, radiative cooling, star and massive black hole formation, as well as feedback from stars and active galactic nuclei. The detailed implementation is discussed in Weinberger et al. (2017) and Pillepich et al. (2018a). The simulation uses a flat Lambda cold dark matter cosmological model with $h = 0.6774$, $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, and $\Omega_{b,0} = 0.0486$, which we adopt for this analysis. TNG follows nine chemical elements: H, He, C, N, O, Ne, Mg, Si, and Fe. Every stellar particle in the simulation inherits the gas elemental abundance

out of which it formed and takes into account chemical enrichment via three distinct stellar evolutionary channels: asymptotic giant branch stars, core-collapse supernovae, and type Ia supernovae. The produced elements are released to the surrounding gas at the end of the stellar lifetime (the detailed algorithm and yield tables used can be found in Pillepich et al. 2018a, sections 2.3.3 and 2.3.4). Most importantly for this work, the dominant enrichment channel of magnesium and oxygen are core-collapse supernovae of massive stars (8–100 $M_{\odot}$), leading to an almost instantaneous enrichment due to their short lifetime. Iron enrichment happens to a large degree via type Ia supernovae, following a more extended delay time distribution that leads to a more gradual enrichment over time. In the gas phase, each cell contains information about the same nine elements, which are treated as passive scalars and advected with the gas flow without impacting it (other than indirectly via the radiative cooling rate, which is metallicity-dependent). This implies that the gas elemental abundance at the position of a star particle is simply given by the respective abundance of the gas cell in which this star particle is located. To convert mass fractions provided by these elemental abundances to relative metallicities, we adopt solar values from Asplund et al. (2009) and atomic weights from Prohaska et al. (2022).

In this work, we are analysing the TNG100-1 simulation, i.e. the highest resolution run of the TNG100 volume. TNG100 refers to the mid-sized simulation box with a volume of 110.7^3 Mpc^3. With such dimensions, the TNG100 box provides a balanced look into both large-scale phenomena such as galaxy clustering and small-scale properties like galaxy structure and gas composition. Using the highest resolution run with 2×1820^3 resolution elements, each galaxy considered in this study contains at least 2×10^4 star particles, sufficient to analyse not only their global properties but also their radial structure.

2.2 Sample selection

We select $z = 2$ galaxies with stellar masses $M_{\star} \geq 10^{10.5}$ $M_{\odot}$, and obtain 1045 objects from the publicly available simulation data (Nelson et al. 2019). For each galaxy we then calculate the gas fraction f_{gas}:

$$f_{\text{gas}} \equiv \frac{M_{\text{gas}}}{M_{\text{gas}} + M_{\star}} \, , \qquad (1)$$

where M_{gas} is the gas mass and $M_{\star}$ is the stellar mass. We consider in particular $f_{\text{gas,1kpc}}$, i.e. the gas fraction within the central 1 kpc. Fig. 1 shows the distribution of $f_{\text{gas,1kpc}}$ for the TNG galaxy population at $z = 2$. The central gas fraction follows a clearly bimodal distribution, which we use to classify galaxies. We define quiescent galaxies as those with $\log f_{\text{gas,1kpc}} \leq -3$ (505 systems at $z = 2$), and as star-forming galaxies those with $\log f_{\text{gas,1kpc}} > -3$ (540 systems).

Our selection of quiescent systems is based on the central gas fraction instead of the more commonly adopted star-formation rate because this choice leads to a greater similarity in radial gas profiles for star-forming galaxies, which will be discussed in Section 4. The galaxy-wide measurements, discussed in Section 3, do not qualitatively change when adopting a selection based on star-formation rate. This is expected, since in TNG there is a close relation between central gas mass and star-formation rate (Nelson et al. 2018; Terrazas et al. 2020).

In the TNG simulation, the central gas fraction is bimodal over the redshift range $0 < z < 4$, and we use the same selection to extract two comparison samples at lower redshifts. Using identical criteria in stellar mass and gas fraction, we select 2452 massive galaxies (of
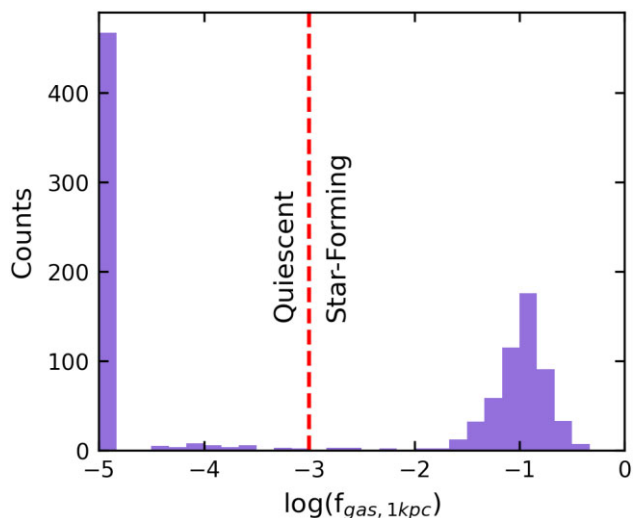
Figure 1. The bimodal distribution of central gas fraction for high-mass $z = 2$ TNG galaxies. Outliers lower than -5 dex are automatically placed in the leftmost visible bin. These tend to have low gas density centres where a single cell of $\sim 10^6\,\mathrm{M}_\odot$ fills the entire inner kpc. The red dotted line at -3 dex marks the cutoff between star-forming (right-hand panel) and quiescent (left-hand panel) galaxies.

which 1618 are quiescent) at intermediate redshift, $z = 0.7$, and 2855 massive galaxies (of which 2423 are quiescent) at $z = 0$.

3 GLOBAL STELLAR ABUNDANCES

For the majority of galaxies at intermediate and high redshift, observational studies can only obtain global (i.e. galaxy-wide) measurements of stellar abundances. In this section, we thus focus on the global stellar abundances of simulated galaxies, which we calculate by taking the luminosity-weighted average of all stellar particles within the effective radius (i.e. the radius containing half of the total stellar mass of the galaxy). We only consider the abundances of magnesium and iron, which are the elements that can be more easily measured in observations.

3.1 Scaling relations

Fig. 2 shows [Mg/H], [Fe/H], and [Mg/Fe] as a function of galaxy mass (left-hand panel) and effective radius (right-hand panel). Each small circle represents an individual TNG galaxy at $z = 2$; quiescent galaxies are shown in red, star-forming galaxies in grey. Red lines represent linear fits to the quiescent TNG galaxies. The TNG simulation makes clear predictions regarding the stellar abundances of massive galaxies at $z = 2$:

(i) The abundances of magnesium and iron are approximately flat as a function of stellar mass. This is analogous to the observed mass–metallicity relation in the local Universe, which features a positive trend for low-mass systems but is mostly flat for $M_\star > 10^{10.5}\,\mathrm{M}_\odot$ (Gallazzi et al. 2005).

(ii) The relative abundance [Mg/Fe] increases slightly with stellar mass; however, all quiescent galaxies are found within a remarkably narrow range of values, with a spread of ~ 0.1 dex.

(iii) Both [Mg/H] and [Fe/H] feature a negative trend with the galactic effective size, with logarithmic slopes of -0.15 and -0.19, respectively. On the other hand, their ratio [Mg/Fe] follows a mildly

positive trend with size (logarithmic slope: 0.04). These trends will be the subject of further investigation in Section 4.

(iv) Star-forming galaxies tend to be slightly more enriched than quiescent galaxies both at fixed mass and at fixed radius, especially in [Fe/H]; they also have lower [Mg/Fe]. This is consistent with their star-formation histories being more extended, thus providing more time for type Ia supernovae to go off and deposit iron in the interstellar medium prior to the next generation of stars being formed.

In Fig. 2, we also include high-redshift observations from Kriek et al. (2016, 2019) and Jafariyazani et al. (2020), shown as large black markers with error bars. All these observational studies use similar types of data and identical spectral fitting techniques to derive stellar abundances. The simulated galaxy population is in broad agreement with the observational data, but the sparse sampling and large uncertainties of the observations prevent a detailed comparison. One notable exception is the ultramassive galaxy observed by Jafariyazani et al. (2020), whose stellar abundance is significantly larger than what we found in the simulation, particularly for [Mg/H]. Moreover, using a different method based on spectral indices, Lonoce et al. (2020) measured a wide range of abundances (spanning from -0.8 to $+0.6$) for a sample of four massive galaxies at high redshift. We note, however, that at these high masses the TNG galaxy sample is incomplete due to limitations in the simulated volume.

Given the difficulty of comparing the TNG predictions to the observations at $z \sim 2$, we also perform the comparison at lower redshift. In Appendix A, we show the TNG abundance relations for the $z = 0.7$ snapshot (which are qualitatively similar to those found at $z = 2$), compared to the abundances measured by Beverage et al. (2021) for galaxies observed at the same redshift. The observed [Mg/H] and [Fe/H] scaling relations are systematically offset compared to the simulations, but feature similar slopes. In particular, the observed abundances of quiescent galaxies appear to follow a flat trend with mass and a negative trend with radius. The anticorrelation between stellar abundances and galaxy sizes has also been pointed out by several observational studies carried out at low and intermediate redshift (McDermid et al. 2015; Barone et al. 2018, 2022; Beverage et al. 2021). The relative abundance [Mg/Fe], on the other hand, does not show appreciable trends in the observations, and features only a mild trend with mass in the simulations.

3.2 Abundance patterns

The relation between [Mg/Fe] and [Fe/H] is often used to probe the time-scale of galaxy formation with the aid of chemical models (e.g. Kriek et al. 2016). In the simplest scenario, when galaxies form rapidly their abundance pattern is mostly set by the material produced by short-lived massive stars, and thus features high [Mg/Fe] and low [Fe/H]. In contrast, in the case when systems form slowly, there is sufficient time for low-mass stars to affect the composition of later stellar generations. As a result, [Mg/H] is low and [Fe/H] is high. We show the abundance patterns of simulated galaxies in Fig. 3; in each panel we colour code the quiescent systems according to the median stellar age, the formation time-scale (defined as the difference between the 90th and 10th percentile of the stellar age distribution), and the effective size, respectively. The first panel also includes the same observations as in Fig. 2. Across all panels, the same fit to the [Mg/Fe] abundance as a function of [Fe/H] is shown in red, highlighting a relatively tight negative correlation. This trend is apparently consistent, at least qualitatively, with the prediction of the simple chemical models described above. However, the colour coding reveals that this anticorrelation between [Mg/Fe] and [Fe/H]
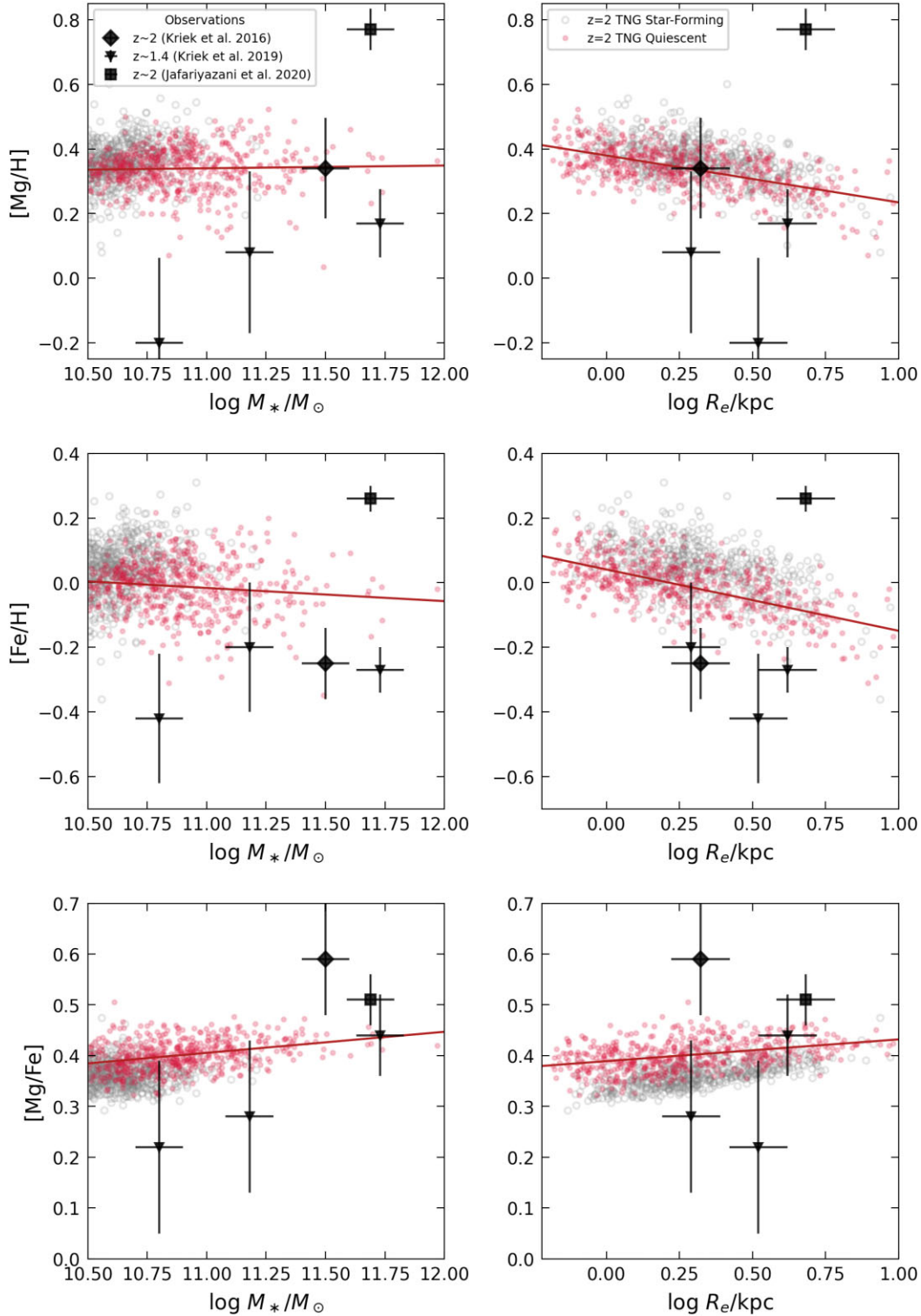
Figure 2. Stellar abundance ratios for high-redshift galaxies as a function of mass (left-hand panel) and effective radius (right-hand panel). TNG galaxies at $z = 2$ are shown as small circles, with quiescent systems in red and star-forming galaxies in grey. In each panel the red line is a linear fit to the simulated quiescent galaxies. Larger, black symbols mark observed galaxies (Kriek et al. 2016, 2019; Jafariyazani et al. 2020). We adopt a stellar mass uncertainty of 0.1 dex for the observations, which is a lower limit on the typical statistical error (Mobasher et al. 2015).

is *not* driven by changes in formation time-scale or median age. The effect of the formation time-scale is mostly vertical, with galaxies that form rapidly being more enriched in [Mg/Fe] at fixed [Fe/H]. Stellar ages have an analogous effect but in the opposite direction,

consistent with the fact that galaxies that form slowly are, on average, younger. The star-forming population is found at the lower [Mg/Fe] envelope of the overall distribution, again consistent with having the youngest stellar ages and longest star-formation time-scales. The
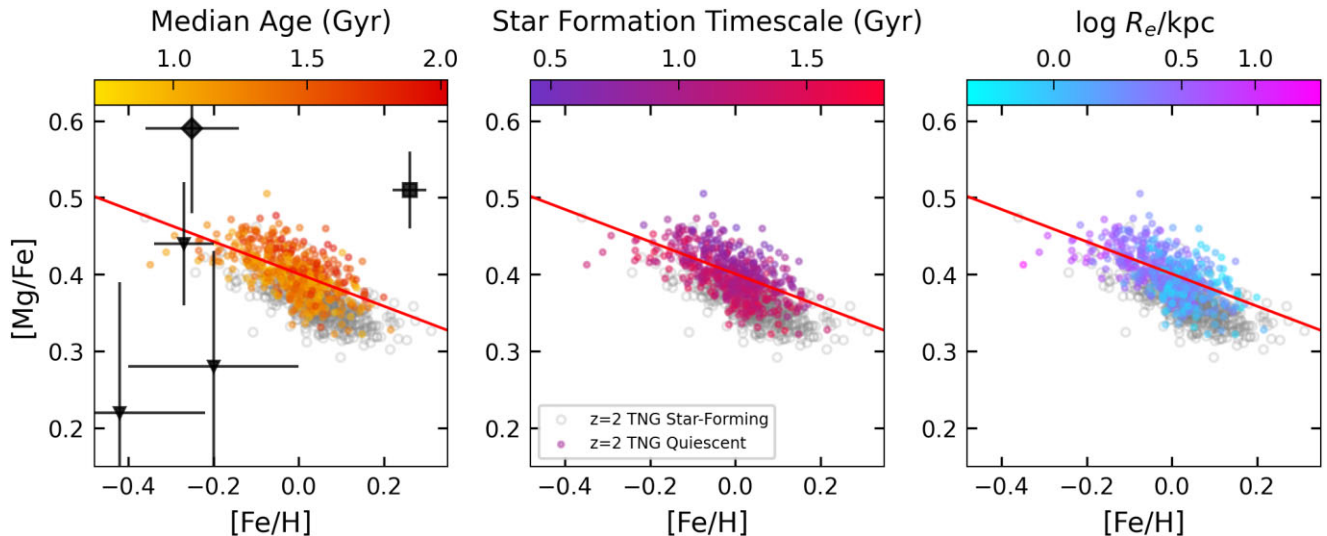
Figure 3. [Mg/Fe] versus [Fe/H] for $z = 2$ simulated galaxies. In each panel, quiescent galaxies are colour-coded according to a different variable: the median stellar age, the star-formation time-scale (both calculated within one effective size), and the effective size. In the left-hand panel we also show the observations at $z \sim 2$ (Kriek et al. 2016, 2019; Jafariyazani et al. 2020), following the legend in Fig. 2. The red line is a linear fit to the quiescent simulated galaxies.

strongest trend along the sequence is, instead, with the effective size: Smaller galaxies have systematically lower [Mg/Fe] and higher [Fe/H]. Thus, the sequence that we see in this diagram is tightly linked to the trends with effective size previously shown in Fig. 2.

In Appendix A, we analyse the abundance patterns for simulated galaxies at $z = 0.7$, and compare them to the available observations.

3.3 Redshift evolution

Next, we explore the galaxy-wide stellar abundances as a function of redshift. While iron is often taken as the reference element for historical reasons, from now on we will focus on the magnesium abundance, which is set by core-collapse supernovae and thus can be easily interpreted in the context of the instantaneous recycling approximation (see e.g. the discussion in Weinberg et al. 2019). Moreover, magnesium traces the total metallicity [Z/H] much more closely than iron, as can be seen by taking the Thomas, Maraston & Bender (2003) relation: [Z/H] = [Fe/H] + 0.94 [Mg/Fe], and rewriting it as [Z/H] = 0.06 [Fe/H] + 0.94 [Mg/H]. Accordingly, we will use 'stellar metallicity' to indicate the magnesium abundance.

Fig. 4 shows [Mg/H] as a function of effective size for simulated galaxies at $z = 2$ (left-hand panel), $z = 0.7$ (central panel), and $z = 0$ (right-hand panel). All plots include the $z = 2$ quiescent trend line for reference. Black lines on the right two panels mark the trends of quiescent galaxies in their respective redshifts. The anticorrelation between [Mg/H] and effective size remains in place from $z = 2$ to $z = 0$, but with a shallower slope and a larger scatter at later cosmic times. Towards low redshift, massive galaxies also grow larger in size, as indicated by the rightward migration of the sample. The star-forming galaxy population tends to have slightly higher [Mg/H] values than the quiescent population for all three redshifts.

The relation between [Mg/Fe] and effective radius is shown in Fig. 5 for different redshifts. This relation becomes shallower and its scatter increases with cosmic time, similar to what is seen for [Mg/H]. Overall, the clear trend with effective radius seen at $z = 2$ is significantly weakened by $z = 0$. Star-forming galaxies tend to have lower [Mg/Fe] values than quiescent galaxies at the same redshift, implying that they are more enhanced in [Fe/H] than in

[Mg/H] relative to the quiescent population. While the average value of [Mg/Fe] in the star-forming population decreases towards lower redshift, they always make up the lower envelope of the overall galaxy population. Finally, Fig. 6 shows that the decrease in average [Mg/Fe] with cosmic time is a mass-independent effect, since the slope of the [Mg/Fe] versus stellar mass relation does not appear to evolve with redshift. Fig. 6 further shows that the lower [Mg/Fe] in the star-forming population is also present at fixed stellar mass.

In conclusion, the scaling relations of simulated galaxies are substantially tighter at high redshift than at $z = 0$. This likely reflects the importance of processes such as mergers, which can mix the stellar populations of different galaxies. Even in the absence of external processes, the abundance evolution of the quiescent galaxy population is affected by quenching: At each redshift, newly formed quiescent galaxies are added to the population, and their abundance follows the abundance distribution of star-forming galaxies at that redshift (see e.g. Choi et al. 2020). As a result of this process, the abundance distribution widens towards lower redshifts. The $z = 2$ relations are thus more 'pristine', and offer a cleaner way to both probe the physics of galaxy formation and test the TNG model using observations.

Finally, we caution that Figs 4–6 show different samples at different redshifts. As a result, the evolution of the scaling relations is due to a combination of physical evolution and a change in the sample composition, making it difficult to develop a simple, meaningful interpretation. For example, Fig. 4 shows that the typical [Mg/H] abundance decreases with cosmic time, at fixed radius. However, this is mostly due to the contribution of metal-poor galaxies that enter the mass selection at low redshift: When we follow the $z = 2$ sample forward in time, we find that the [Mg/H] versus effective size relation *increases* with time. Moreover, as each individual galaxy becomes bigger due to satellite accretion, the abundance measurement probes a different physical region of the galaxy, further complicating the interpretation. This effect is particularly important when interpreting the star-forming population, which can be considered as being in a transition state (entering the selection via stellar mass growth and leaving it via quenching, which is a one-way process for most galaxies).
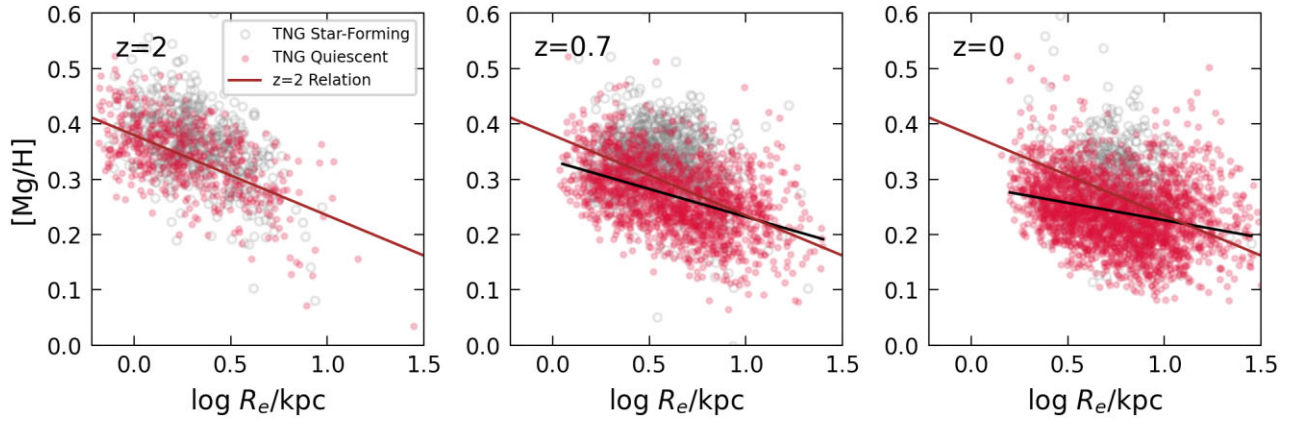
Figure 4. Stellar [Mg/H] versus effective size for simulated galaxies, in three redshift bins. The red line, seen in all panels, is a linear fit to the $z = 2$ quiescent systems. The black lines in the last two panels are linear fits to the quiescent galaxies at their respective redshift.
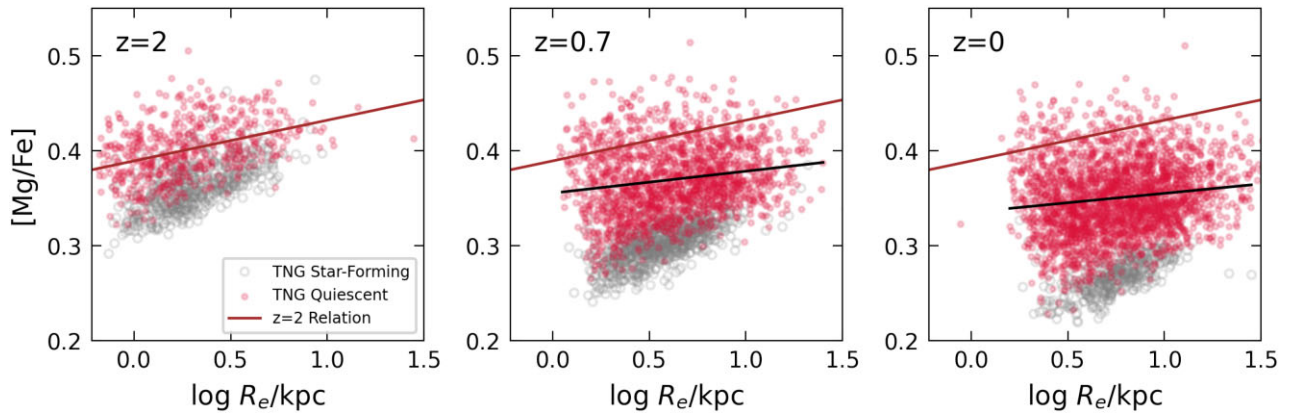
Figure 5. Stellar [Mg/Fe] versus effective size for simulated galaxies, in three redshift bins. The red line, seen in all panels, is a linear fit to the $z = 2$ quiescent systems. The black lines in the last two panels are linear fits to the quiescent galaxies at their respective redshift.
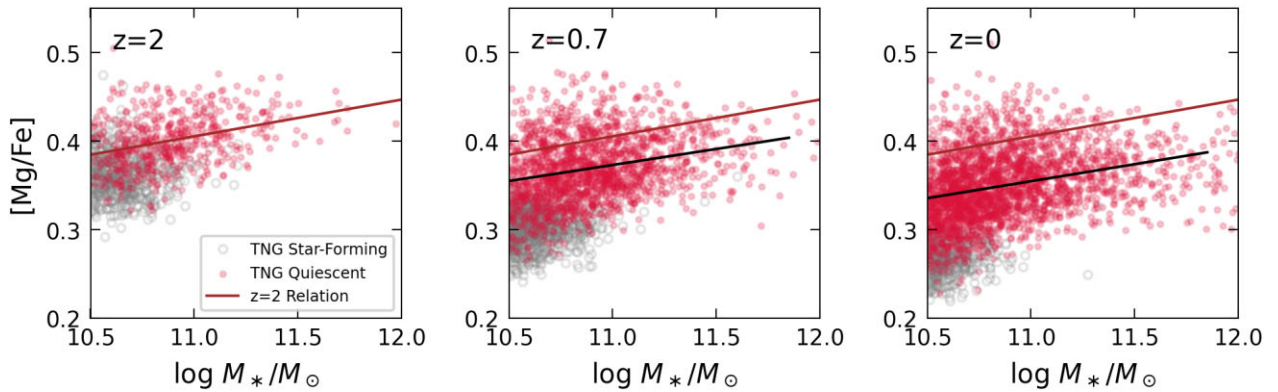
Figure 6. Stellar [Mg/Fe] versus stellar mass for simulated galaxies, in three redshift bins. The red line, seen in all panels, is a linear fit to the $z = 2$ quiescent systems. The black lines in the last two panels are linear fits to the quiescent galaxies at their respective redshift.

3.4 Aperture dependence

One crucial aspect of stellar abundances, that must be carefully considered when comparing simulations with observations, is the aperture over which the abundances are measured. We test for this effect in Fig. 7, where we show the abundances of simulated galaxies calculated using two different apertures. The red symbols show the luminosity-weighted average abundance of all stars within the central

effective radius (which is the aperture used to generate all previous figures), while the black symbols show the luminosity-weighted average abundance of all stars within the central kpc. For simplicity we restrict ourselves to the quiescent population at $z = 2$.

The choice of the aperture is clearly important for the resulting scaling relations: The left-hand panels of Fig. 7 show that the abundance in the central kpc features a significant increase with stellar mass, as opposed to the flat relation that we find when using
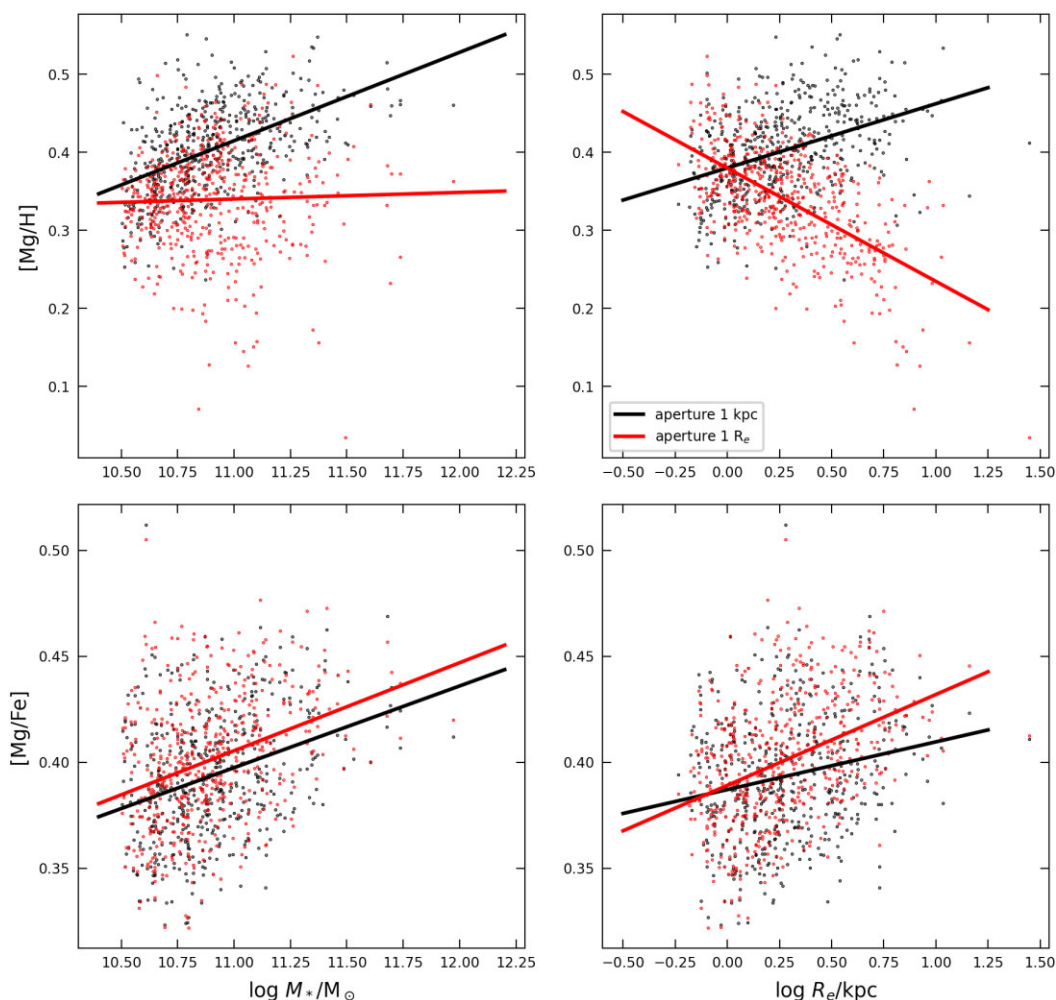
Figure 7. Abundance ratio measured in different apertures versus stellar mass (left-hand panel) and versus effective radius (right-hand panel) for quiescent $z = 2$ galaxies. Red points show the abundances calculated within 1 effective size (like in the previous figures), while black points show the abundances calculated within a fixed aperture of 1 kpc. Solid lines represent linear fits to the data.

the abundances calculated within the central effective size. Most strikingly, the right-hand panels of Fig. 7 show that the abundance trend with radius is flipped, switching from a negative to a positive slope when changing the aperture from 1 R_e to 1 kpc. We note that the trend for a measurement within a fixed aperture of 30 kpc (not shown here) yields similar slopes than when using the effective radius. The relative abundance [Mg/Fe], on the other hand, does not depend strongly on the aperture used.

4 RADIAL PROFILES

In this section, we move beyond galaxy-wide measurements and explore the spatial distribution of metal abundances within simulated galaxies. This is necessary in order to understand the aperture dependence of the scaling relations, and to investigate the physical origin of the abundance relations, including the connection between stellar and gas-phase metallicity.

4.1 Stellar abundance profiles

Fig. 8 shows the stellar abundance profiles of 100 randomly chosen simulated quiescent galaxies at $z = 2$ as a function of radius, expressed in units of kpc (left-hand panel) and in units of the

effective radius (right-hand panel). The top panels show [Mg/H] and the bottom panels show [Mg/Fe]. The [Mg/H] profiles are steeply declining, spanning a dynamic range of roughly 1 dex. The [Mg/Fe] profiles, in contrast, are rising and shallow, spanning only ∼0.1 dex. The scatter in the bottom panels appears to be larger because of the different scale, while in reality it is much smaller. Interestingly, the scatter in the [Mg/H] profiles in the inner parts is smaller when plotted against radius in kpc instead of in units of the effective size. This indicates that the process setting the abundance profile is different from the process responsible for setting galaxy sizes. Note, however, that in the outskirts (outside of ∼3 kpc or ∼1 effective radius) the opposite seems to be the case.

The steepness of the [Mg/H] profiles, together with the comparably low scatter of the profiles as a function of physical radius, explain the aperture dependence of the scaling relations discussed in Section 3.4. When considering the abundance within the effective size, small galaxies only 'probe' the central, enriched part of this quasi-universal profile, while larger galaxies 'probe' a substantial number of distant and metal-poor stars, bringing down the luminosity-weighted abundance average by 0.1–0.2 dex. On the other hand, measuring the enrichment in an aperture of 1 kpc only probes the galactic centre, which differs from the galaxy-wide measurement in the case of more extended galaxies, thus altering the slope of the [Mg/H]
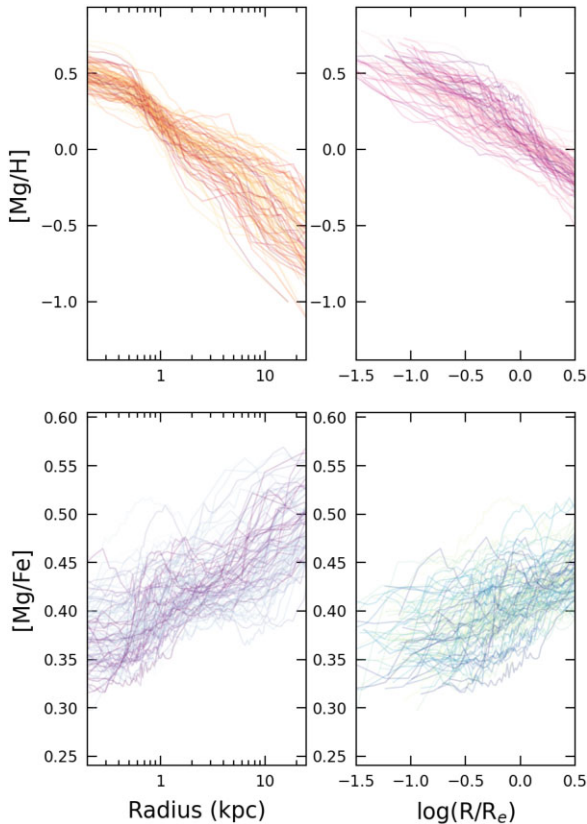
Figure 8. Radial profiles of stellar [Mg/H] (top panel) and [Mg/Fe] (bottom panel) for 100 random quiescent TNG galaxies at $z = 2$. The profiles are shown both as a function of physical radius in kpc (left-hand panel) and as a function of radius normalized by the galaxy effective size (right-hand panel). Note that the bottom panels span a much smaller dynamic range.

versus effective radius relation. The [Mg/Fe] ratio suffers less from this effect since the gradient is less steep.

What remains to be explored is the origin of the steepness, the small scatter, and the seemingly universal stellar metallicity profile in the inner regions of $z = 2$ galaxies, for which we have to turn towards the connection between gas enrichment and stellar metallicity.

4.2 Gas versus stellar metallicity

Since stars inherit the metallicity of the gas from which they form, we need to explore the gas-phase abundance if we want to understand the origin of the stellar abundance profiles. However, quiescent galaxies are devoid of gas (at least in the centre) by selection; we thus turn to the massive, $z = 2$ star-forming galaxies as examples of progenitor galaxies, and study their gas properties. From now on, we focus on the abundance profiles as a function of physical radius in units of kpc, in light of the lower scatter. Moreover, to remain as close as possible to the observations, we will consider the abundance of oxygen instead of magnesium. While they are both α elements, the oxygen abundance is much easier to measure for the gas, particularly at high redshift, and is often taken as the 'gas metallicity'. We checked that the stellar [Mg/H] and [O/H] profiles are virtually identical so that the conclusions we draw from analysing the oxygen abundance can also be applied to magnesium.

We compare gas and stellar abundances for 100 random star-forming galaxies at $z = 2$ in Fig. 9. The first panel shows the stacked [O/H] abundance profile for all 100 star-forming galaxies: The colour

coding represents the number of stellar particles (belonging to any galaxy) that fall within each two-dimensional bin. The middle panel is coloured according to the median age of star particles in each bin, ranging from 0 to 2.5 Gyr. Black contour lines overplot the particle density as seen in the first panel. The overall abundance profile appears to be due to the superposition of stellar populations of different ages, each featuring an abundance profile with roughly the same slope but varying normalization from metal-poor for the oldest stars to metal-rich for the youngest ones. The third panel shows the gas-phase metallicity profile, which is very similar to the distribution of the youngest stars seen in the second panel. This confirms that the slope of the stellar abundance profile is primarily set by the slope of the gas-phase abundance profile. However, the stellar abundance profile is also modified by different formation times of stars at different radii and its scatter at fixed radius, which, due to a time evolving gas metallicity profile, causes a larger scatter in stellar metallicities at fixed radius as well as additional features in the profile. Therefore, the steepness of the metallicity gradients in quiescent galaxies is a remnant of steep gas-phase metallicity gradients during their star-forming phase (assuming little radial mixing has occurred, which is a reasonable assumption at $z = 2$). The low galaxy-to-galaxy variation of stellar metallicity further indicates that the gas metallicity profiles during the star-forming phase were very similar as well.

4.3 Gas fraction and effective yield

To further understand the origin of the gas metallicity profile, in Fig. 10 we explore the properties of the gas in the sample of simulated star-forming galaxies. The first panel shows the radial profiles of the gas consumption time-scale, defined as the ratio between the local gas mass density and the local star-formation rate density. There is a clear spatial variation, with the galaxy centres being efficient at forming stars (with time-scales of a few hundred Myr), and the galaxy outskirts being inefficient, taking about 3 Gyr to transform the gas into stars – a time comparable to the age of the Universe at $z = 2$. As a consequence, the gas fraction (shown in the second panel) is only a few per cent in the centre, where most of the gas has already been transformed into stars (or removed by galactic outflows), but approaches 100 per cent in the outskirts.

The third panel of Fig. 10 shows the gas-phase abundance profile, which is clearly anticorrelated with both the consumption time-scale and the gas fraction: Regions where most of the gas has been transformed into stars feature a high metallicity, while regions where gas is transformed very slowly are metal-poor. This behaviour is in qualitative agreement with the expectation of the simplest model of chemical evolution: the closed-box system. In this model, an initial mass of metal-poor gas is transformed into stars without further inflows or outflows. The model is characterized by a simple relation between the gas metallicity and the gas fraction, which depends only on the effective yield, defined as

$$y_{\mathrm{eff}} \equiv \frac{Z_{\mathrm{gas}}}{\ln(1/f_{\mathrm{gas}})} \, , \tag{2}$$

where Z_{gas} is the fraction of the total baryonic mass contributed by a specific metal. In our case, we consider oxygen, so $Z_{\mathrm{gas}} = 10^{\mathrm{[O/H]}} \cdot (M_O/M_{\mathrm{tot}})_\odot$. We show the radial profile of the effective yield in the last panel: for most galaxies, this quantity remains approximately constant at around 10^{-4}, out to ~10 kpc. This is a remarkable result: while the gas fraction and the gas abundance vary by more than an order of magnitude along the Galactocentric radius, the effective yield, which is a simple combination of these two quantities, is roughly constant. Such result suggests that galaxies are, indeed,
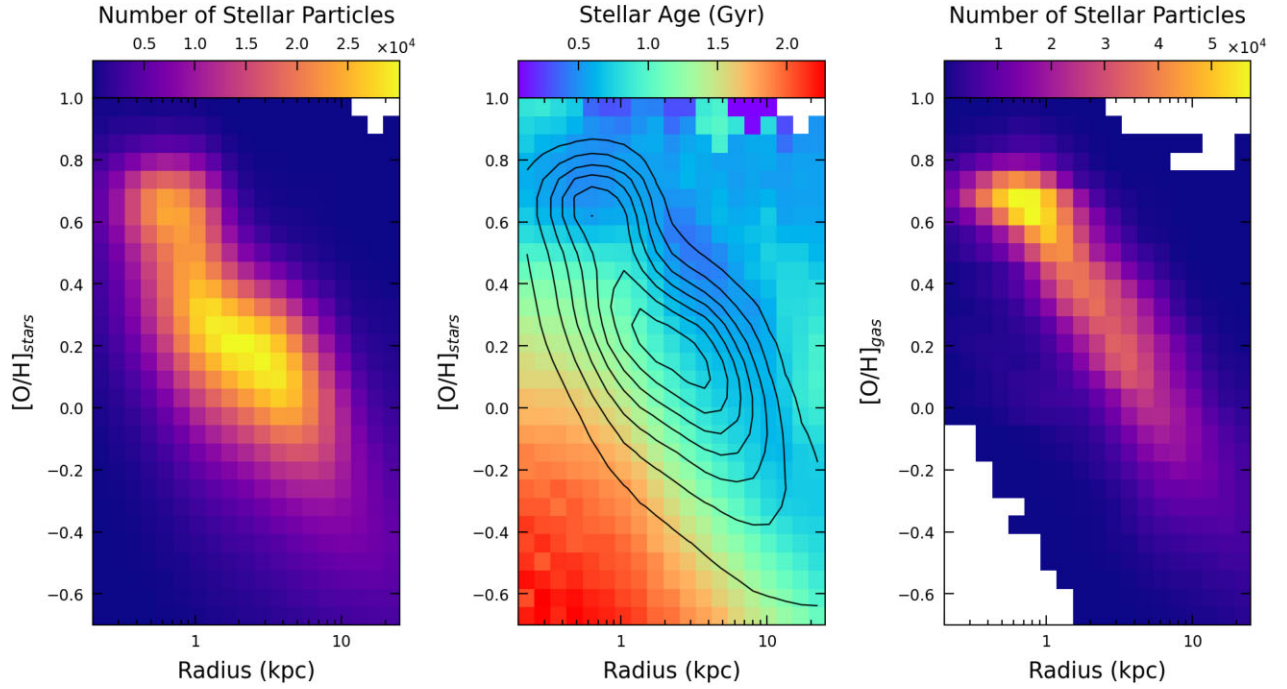
Figure 9. Metallicity versus radius for 100 random star-forming TNG galaxies at $z = 2$. Left-hand panel: The colour coding shows the number of stellar particles (belonging to any galaxy) in each bin of stellar [O/H] and Galactocentric radius. Central panel: The colour coding shows the median stellar age in each bin; the number of stellar particles shown in the previous panel is overlaid in black. Right-hand panel: Same as the first panel, but for gas-phase [O/H].
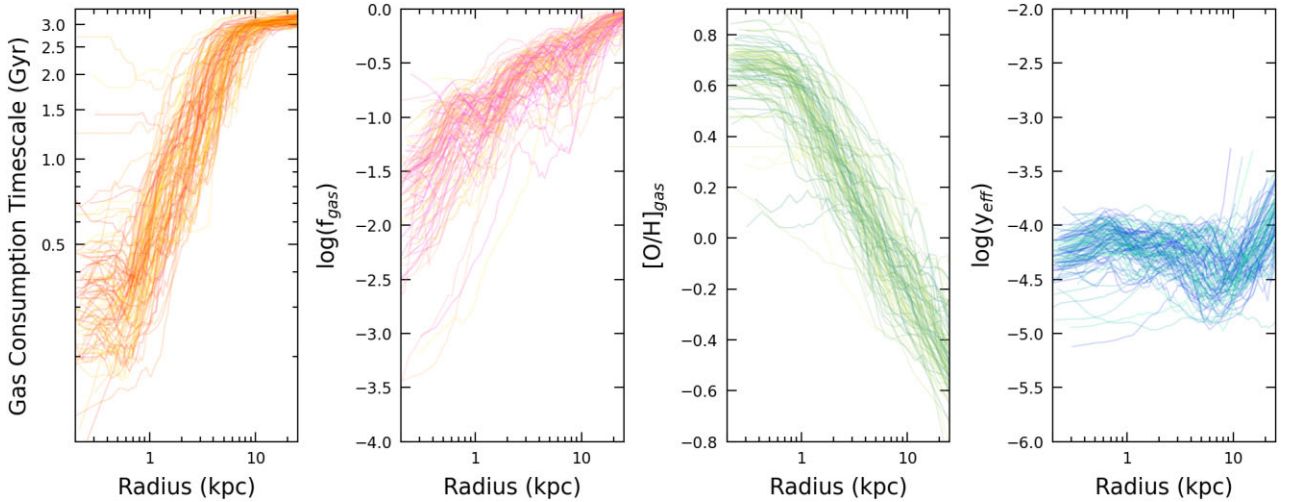
Figure 10. Radial profiles of gas properties for the same random 100 star-forming TNG galaxies at $z = 2$ shown in Fig. 9. From left to right, the panels show the radial profiles for the gas consumption time-scale (i.e. local gas mass divided by local star-formation rate), gas fraction, gas-phase oxygen abundance, and effective yield as defined in equation (2).

similar to closed-box systems *at fixed radius*. The local closed-boxed systems at two different radii have the same effective yield, but due to their different star-formation efficiencies they are in two different phases of evolution, characterized by different values of gas fraction and metallicity.

4.4 Closed-box model

Motivated by the approximate flatness of the effective yield profile shown in Fig. 10, we construct a toy model in which each radial shell of a galaxy is a closed-box system. We also assume the instantaneous

recycling approximation (that is, there is no delay between star-formation and the return of enriched material to the interstellar medium), which is appropriate for α elements such as magnesium and oxygen. In this form, the closed-box system has been extensively studied, and analytical solutions exist (see Matteucci 2012, for a review).

We test our toy model by exploring the distribution of stellar abundances. The closed-box model predicts a simple function for $S(Z)$, the fraction of stars with metallicity lower than Z:

$$S(Z) = \frac{1 - (f_{\text{gas}})^{Z/Z_{\text{final}}}}{1 - f_{\text{gas}}} \ . \tag{3}$$
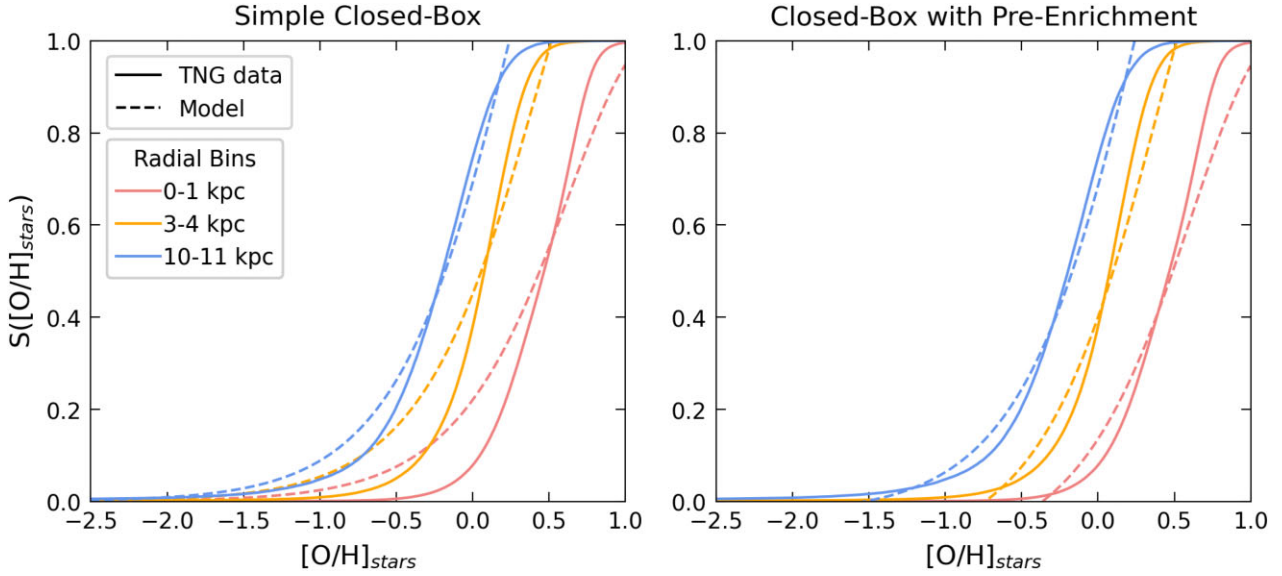
Figure 11. The cumulative distribution function, describing the fraction of stars below a given value of [O/H]. In both panels, the solid lines show the distribution function measured for 100 star-forming TNG galaxies at $z = 2$ in three different radial bins. The dashed lines show the best fit for two toy models: the simple closed box in the left-hand panel, and the pre-enriched closed box in the right-hand panel. Note the improved fit at low metallicities once pre-enrichment is included.

Table 1. Parameter values for the toy model described in Section 4.4. f_{gas} and [O/H]$_{gas}$ are the median values measured in each radial bin. [O/H]$_{final}$ is the best fit to the simple closed-box model given in equation (3), while [O/H]$_{initial}$ is the pre-enrichment found to provide the best fit when using the pre-enriched closed-box model given in equation (4). We used $Z \propto 10^{[O/H]}$.

Radial bin (kpc)	f_{gas}	[O/H]$_{gas}$	[O/H]$_{final}$	[O/H]$_{initial}$
[0, 1]	0.045	0.643	1.119	−0.364
[3, 4]	0.271	0.192	0.516	−0.739
[10. 11]	0.370	− 0.124	0.240	−1.524

The metallicity distribution of a closed-box system thus depends only on the current gas fraction f_{gas} and the final stellar metallicity Z_{final}, which coincides with the current gas-phase metallicity and represents the maximum possible value for the stellar metallicity. In the left-hand panel of Fig. 11, we compare this prediction to the stellar oxygen abundances measured for star-forming TNG galaxies in three radial bins, each with a width of 1 kpc. For each bin we adopt the median value of the measured f_{gas}, and then apply a least-squares fit to find Z_{final}. Fit values are listed in Table 1; note that in the table and figures we convert all mass fractions Z to the corresponding logarithmic abundances [O/H]. The model is able to match the median [O/H] in each radial bin, and broadly reproduces the shape of the cumulative distribution of oxygen abundance. This means that the toy model can explain the overall trend of the stellar metallicity profiles. However, we note two important discrepancies. First, the fits return values of [O/H]$_{final}$ that are 0.3–0.4 dex larger than the gas-phase abundance measured in the same bins (see Table 1). This suggests that the gas metallicity was higher at some point in the past, which is impossible in a closed-box system. Second, our simple model produces too many stars with low metallicity compared to the TNG simulation (see Fig. 11, left-hand panel). This is analogous to the discrepancy between the closed-box model prediction and observations of stellar metallicities in the Milky Way (e.g. Tinsley 1980). One simple way to avoid this discrepancy is to assume that
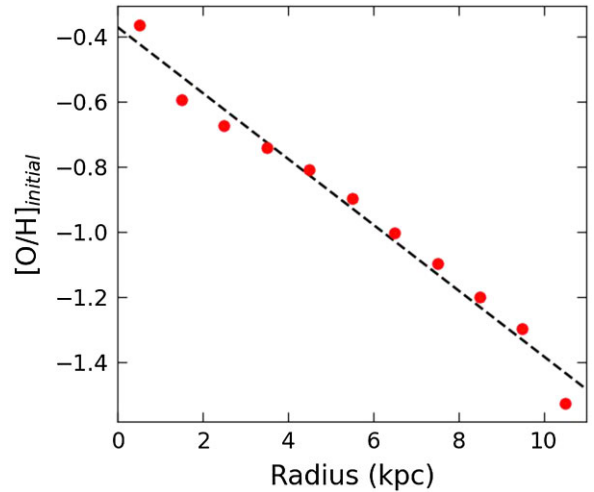
Figure 12. The best-fitting pre-enrichment abundance for 100 star-forming TNG galaxies at $z = 2$, based on equation (4). Each red point represents the centre of a 1-kpc radial bin. The linear fit, shown in black, has a slope of approximately –0.10 dex per kpc.

the initial gas is pre-enriched, so that even the earliest generations of stars are relatively metal-rich. In this case, the expected distribution of stellar metallicity becomes

$$S(Z) = \frac{1 - (f_{gas})^{(Z - Z_{initial})/(Z_{final} - Z_{initial})}}{1 - f_{gas}} , \qquad (4)$$

where $Z_{initial}$ is the metallicity of the gas at the beginning of the closed-box evolution. Using this equation, and adopting the same f_{gas} and Z_{final} found above, we fit the TNG metallicities for $Z_{initial}$ in each radial bin and show the results in the right-hand panel of Fig. 11. Clearly, this model improves our prediction of the TNG stellar abundances. We repeat the fitting procedure in several 1-kpc radial bins, and plot the resulting profile of the initial gas abundance in Fig. 12. The pre-enrichment value is about two-fifths of the solar metallicity

(−0.4 dex) in the galaxy centre and declines with a roughly constant gradient of −0.10 dex per kpc.

In conclusion, we find that the *gas-phase* metallicity as a function of radius can be modelled by a series of closed-box shells with identical effective yield. However, such a model is too simplistic to describe the distribution of *stellar* metallicity, which can be explained with closed-box models only if one allows a radially varying pre-enrichment. The pre-enrichment model is also very simplistic, and represents only one of the several possible solutions. Given the discrepancies arising from the use of a closed box model to describe the TNG data, we conclude that gas inflows and outflows are likely to play an important role in determining the stellar abundance profiles in TNG.

5 DISCUSSION

5.1 The gas-phase abundance profile

The panels in Fig. 10 provide a clear narrative of star formation and chemical enrichment in $z = 2$ star-forming galaxies. The short gas consumption time-scale in the centre of galaxies means that these regions are more efficient at forming stars. Gas is consumed more rapidly and is depleted, as evinced by the lower gas fraction in the second panel. Thus, the gas that is left is highly enriched, as shown in the third panel. In this rather simplistic scenario, *the gas metallicity profile is directly set by the gas fraction profile*, in agreement with the prediction of the closed-box model. Outside factors like outflows or mergers, which are definitely present in the TNG simulations, do not cause a substantial change in the metallicity profile. The final panel, showing that the effective yield is remarkably flat out to ∼10 kpc, confirms our simple interpretation: Each radial bin can be considered as an independent closed-box system at a different stage of evolution, which is set by the star-formation efficiency at that location. Both real and simulated galaxies are subject to substantial inflows and outflows of gas, and it is therefore surprising that such a simple model is able to capture the gas-phase metallicity profile. However, while gas flows are able to temporarily increase or decrease the metallicity, any subsequent star formation will rapidly bring the system back to the closed-box prediction, as shown in detail by Dalcanton (2007).

The importance of gas fraction as the primary driver of gas-phase metallicity has been highlighted by several studies. For example, Torrey et al. (2019) showed that in TNG, the mass–metallicity relation is explained by the simple fact that low-mass galaxies have a higher gas fraction, and therefore a lower gas-phase metallicity compared to massive systems. Other types of simulations, including zoom-in simulations, provide similar pictures (e.g. Ma et al. 2016). These studies were mostly concerned with the interpretation of galaxy-wide metallicities, rather than radial profiles: In this respect, our approach is more similar to that of Belfiore et al. (2019), who analyse the observed gas metallicity gradients of local galaxies using a simple model consisting of independent radial bins. They adopt the 'bath-tub' model rather than a closed box, but reach similar conclusions: The abundance profile is driven by the radial variation of the star-formation efficiency.

5.2 The stellar abundance profile

Stellar populations are more difficult to model because they feature, at each location, a distribution of abundances tracing the entire enrichment history, as opposed to the gas abundance which is a single value and mostly traces the current properties of the system. Indeed, when applying the closed-box model to the stellar abundance distribution we find inconsistent results. First, at each radial bin we find a fraction of stars with higher metallicity than the gas, which is not possible for a true closed-box system. Second, the model prediction systematically overproduces metal-poor stars compared to the TNG data. We were able to improve the agreement by introducing a pre-enrichment that strongly declines with radius. At face value, this could indicate that the gas reaching the galaxy centres has already been enriched by star-forming regions, while the outskirts are fed by more pristine gas. However, this is an overly simplistic interpretation because the effect of inflows, outflows, and formation time-scales leave a strong imprint on the stellar abundances and cannot be neglected.

A more realistic model of the chemical history would also need to explain the mildly rising [Mg/Fe] profile that characterizes massive galaxies in TNG (see Fig. 8). Using the Versatile Integrator for Chemical Evolution (Johnson & Weinberg 2020), we were able to produce a rising [Mg/Fe] profile by changing the accretion history as a function of radius. In particular, the inflow in the central regions need to start earlier and last longer compared to the inflow in the outskirts. However, in order to also reproduce the gas fraction, stellar [Mg/H], and gas-phase [O/H] profiles, we need to assume a declining profile for the pre-enrichment value, and a rising profile for the mass loading factor. A detailed investigation of chemical evolution models is beyond the scope of this work; however, such preliminary exploration confirms that several physical processes must be considered for a proper understanding of the stellar abundance profiles.

Given the complexity of the processes involved, it is remarkable that the stellar abundance profiles are highly similar when plotted as a function of radius in physical units (i.e. kpc), with a scatter of only ∼0.1 dex. Moreover, the scatter among the profiles is *larger* when normalizing the radial coordinate to the effective size of each galaxy, as is commonly done (see Fig. 8). This fact, which has important consequences for galaxy-wide measurements as we discuss ahead, implies that the abundance profiles do not scale with galaxy sizes but are more or less universal, for massive galaxies at $z = 2$. To first approximation, the stellar abundance profile is set by the past gas-phase abundance profile, which also follows a roughly universal law, as shown in Fig. 10. Older stars being less enriched does not substantially change the slope of the stellar abundance profile. We thus conclude that, to the first order, the stellar abundance profiles in TNG are ultimately driven by the gas fraction profiles. Metal removal via outflows can certainly play a role, since the closed-box model is inadequate to describe the detailed abundance distributions of stars; however, this effect is not primarily responsible for the overall decline in stellar [Mg/H] by ∼0.8 dex over 10 kpc. This is in agreement with the results of Torrey et al. (2019), who found that in TNG, the most massive galaxies are *less* efficient at retaining their metals, which is the opposite behaviour of what is commonly believed.

5.3 Interpreting galaxy-wide abundance measurements

In addition to exploring the physical processes that determine the stellar abundances of massive galaxies, the main goal of this work is to provide a framework for the interpretation of current and future measurements. Stellar abundance measurements are challenging at high redshift, and are often performed on the spatially integrated spectrum of the whole galaxy. Given the steeply declining abundance profiles we find in TNG, our first conclusion is that assessing the physical size probed by the spectroscopic data is paramount. As we discussed in Section 3.4, such aperture effects can have a major impact on the scaling relations. If one assumes that the spectra are

always extracted from the region within R_e, then an anticorrelation between stellar abundance and galaxy size naturally emerges: For small galaxies, R_e probes only the central part of the profile, where the stars are more metal-rich, while for large galaxies R_e is able to probe further into the outskirts of the profile, yielding a lower average abundance. This trend, which we find in the TNG population, is also detected in observations at intermediate redshift (see Appendix A). Our analysis makes a clear prediction: If one could measure the stellar abundances within a fixed physical aperture (e.g. 1 kpc), this trend would disappear or even flip, as shown in Fig. 7.

Ideally, our predictions should be tested at high redshift, but current measurements at $z \sim 2$ suffer from small sample sizes and large uncertainties. The two galaxies with the most accurate abundance measurements, one from Kriek et al. (2016) and the other from Jafariyazani et al. (2020), have similar masses and redshifts but their [Mg/H] abundances differ by 0.4 dex. On the other hand, the [Mg/Fe] relative abundances are consistent within the error bars. This surprising result may be explained if the measurements for the two galaxies were taken at different physical apertures. In particular, the galaxy studied by Jafariyazani et al. (2020) is strongly lensed, and the spectroscopic observations are therefore able to probe the very central part of the abundance profile, yielding an elevated [Mg/H] measurement. Even so, this aperture effect does not have a strong impact on the measurement of [Mg/Fe], which is predicted to have a nearly flat profile. This interpretation is supported by the stellar abundance gradients measured by Jafariyazani et al. (2020), which are the only spatially resolved measurements currently available at high redshift. The observed [Fe/H] declines with a gradient of about -0.040 ± 0.028 dex per kpc, while [Mg/Fe] is approximately flat with radius: Both measurements are in broad agreement with the typical profiles found in TNG galaxies.

Further observations at $z \sim 2$ are required for a robust test of the TNG predictions. With the advent of *JWST*, stellar abundance measurements will become more accessible at high redshift, and in some cases the spatial profiles will also be measured with high precision. Given that TNG produces remarkably tight abundance profiles, it will be easy to test this model. If the TNG predictions are correct, one consequence is that the measured abundance pattern cannot be used to constrain the formation time-scale of massive galaxies at $z \sim 2$ because the intrinsic differences among galaxies, after correcting for aperture effects, are extremely small.

5.4 Caveats

Our results are necessarily tied to the specific physical models implemented in the TNG simulations. For example, the FIRE simulations, which feature a much smaller galaxy sample but higher resolution, obtain very diverse metal gradients at $z = 2$ (although mostly at lower stellar masses, see Ma et al. 2017). This key difference is likely due to the burstiness of the FIRE star-formation model, and may lead to different scaling relations for the stellar abundances. Moreover, the absolute value of the metal abundances measured in TNG are highly dependent on the assumed stellar yields, which are still poorly understood. A preliminary comparison of TNG with the stellar abundances observed in local galaxies suggests that the magnesium yields implemented in the simulation need to be revised (Naiman et al. 2018). Nonetheless, relative measurements such as radial trends should be substantially more robust than the absolute scale of the abundance values.

Finally, we caution that a proper comparison with the observations should account for biases introduced by the measurement methods. It has been shown that the procedure employed to derive stellar abundances from the simulations (such as the way stellar particles are combined, and whether mass- or luminosity-weighting is adopted) and from the observations (e.g. the choice of stellar libraries or the method used to fit templates to the observed spectra) have potentially large effects on the stellar abundances (Guidi, Scannapieco & Walcher 2015; Guidi et al. 2016; Gebek & Matthee 2022). Given that our focus is on the high-redshift Universe, where current measurements are scarce, we do not attempt a realistic comparison between simulations and observations. This effort will be warranted, however, once the sample of available measurements is substantially increased by new observations with *JWST*.

6 SUMMARY AND CONCLUSIONS

Analysing the population of massive ($\log M_*/M_\odot > 10.5$) galaxies at $z = 2$ in the TNG simulations, we investigated the relations between stellar abundances and other global properties such as stellar mass and effective size; the radial dependence of stellar abundances; and the physical origin for the emerging relations. Here, we summarize our findings:

(i) Stellar abundances ([Mg/H] and [Fe/H]) feature a flat relation with stellar mass and an anticorrelation with effective size. The [Mg/Fe] relative abundance has a slightly positive trend with mass and size. Current observations at $z = 2$ are in broad agreement with the TNG predictions, but the small sample size and large uncertainties prevent a meaningful comparison. At lower redshift, these relations become progressively flatter and less tight. Galaxies observed at $z = 0.7$ follow trends that are similar to those found in the simulations.

(ii) Galaxy sizes play an important role in the stellar abundance relations. For example, they drive the difference in [Mg/Fe] between galaxies, while it is commonly assumed that this relative abundance mainly tracks galaxy formation time-scales. Moreover, the aperture used in determining the stellar abundance is critical: If the abundances are measured within the central kpc instead of within R_e, the trend between [Mg/H] and radius is inverted.

(iii) Massive galaxies feature a nearly universal [Mg/H] radial profile, which is steeply declining and does not scale with galaxy sizes. This explains the abundance scaling relations and aperture effects: Smaller galaxies appear metal-rich simply because most of the stars within R_e are found in the inner part of the abundance profile; on the other hand, larger galaxies have more stars within R_e that are found in the outer parts of the profile, where the abundances are low.

(iv) Gas-phase metallicity also follows a nearly universal, steeply declining profile. We develop a toy model in which each radial bin is treated as an independent closed-box system: Since the effective yield is roughly constant with radius, we can easily interpret the gas abundance profile as solely determined by the gas fraction profile. In practice, at the centre of galaxies the gas density is high and gas can be efficiently transformed into stars. This leaves a low gas fraction, which in a closed-box model gives rise to a high metallicity. In the outskirts, on the other hand, star formation proceeds slowly and leaves large gas fractions, resulting in a low gas metallicity.

(v) The stellar abundance profile is set by the gas abundance profile at the time when most stars were formed. To first approximation, then, the stellar abundance profile is also determined by the gas fraction profile. However, our simple closed-box model fails to reproduce the detailed distribution of stellar abundances at a fixed radius. We are able to improve the model by allowing for a radially varying pre-enrichment, but we conclude that a substantially more complex model must be developed in order to explain the stellar abundances in detail.

Perhaps the most important result from our analysis is the discovery of a nearly universal profile for both stellar and gas abundances in simulated massive galaxies at $z = 2$. This sheds light on the physical processes that drive the chemical enrichment of galaxies. At the most fundamental level is the star-formation efficiency, which in the TNG model is simply determined by the gas density. This is what ultimately drives the abundance profile: The central regions are faster at transforming gas into stars and therefore their chemical enrichment is accelerated, resulting in higher levels of gas and stellar abundance. Other processes, such as outflows, likely play some role in determining the detailed distribution of chemical abundances, but are not responsible for the steeply declining abundance profiles.

The existence of a nearly universal profile also has important implications for the interpretation of galaxy-wide abundance measurements. In particular, we showed that the stellar abundances are strongly dependent on the size of the aperture used for the measurement. At the same time, this effect provides a robust test for the simulations, because it produces a relatively tight relation between the stellar abundance measured within R_e and the galaxy effective size.

Our conclusions are, naturally, only valid as long as the physical model employed by the TNG simulations is valid. The extent to which our results apply to the real Universe can only be assessed through a comparison with observations. Since there are intervening processes at lower redshift, such as quenching and galaxy mergers, that tend to dilute the chemical profiles, observations at high redshift are required. The near-infrared capabilities and exquisite sensitivity of *JWST* will soon make this type of measurements possible for representative samples of high-redshift galaxies, leading to conclusive tests of the theoretical predictions presented in this work.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data underlying this article were accessed from the IllustrisTNG simulation website accessible under https://www.tng-project.org/data/. The derived data generated in this research will be shared on reasonable request to the corresponding author.

REFERENCES

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Bahé Y. M., Schaye J., Crain R. A., McCarthy I. G., Bower R. G., Theuns T., McGee S. L., Trayford J. W., 2017, MNRAS, 464, 508
Barone T. M. et al., 2018, ApJ, 856, 64
Barone T. M. et al., 2022, MNRAS, 512, 3828
Belfiore F., Vincenzo F., Maiolino R., Matteucci F., 2019, MNRAS, 487, 456
Beverage A. G., Kriek M., Conroy C., Bezanson R., Franx M., van der Wel A., 2021, ApJ, 917, L1
Borghi N., Moresco M., Cimatti A., Huchet A., Quai S., Pozzetti L., 2022, ApJ, 927, 164
Carnall A. C. et al., 2022, ApJ, 929, 131
Choi E., Brennan R., Somerville R. S., Ostriker J. P., Hirschmann M., Naab T., 2020, ApJ, 904, 8
Christensen C. R., Davé R., Brooks A., Quinn T., Shen S., 2018, ApJ, 867, 142
Dalcanton J. J., 2007, ApJ, 658, 941
De Rossi M. E., Theuns T., Font A. S., McCarthy I. G., 2015, MNRAS, 452, 486
Finlator K., Davé R., 2008, MNRAS, 385, 2181
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Gebek A., Matthee J., 2022, ApJ, 924, 73
Genel S. et al., 2014, MNRAS, 445, 175
González Delgado R. M. et al., 2015, A&A, 581, A103
Grand R. J. J. et al., 2018, MNRAS, 474, 3629
Grand R. J. J. et al., 2019, MNRAS, 490, 4786
Guidi G., Scannapieco C., Walcher C. J., 2015, MNRAS, 454, 2381
Guidi G., Scannapieco C., Walcher J., Gallazzi A., 2016, MNRAS, 462, 2046
Hemler Z. S. et al., 2021, MNRAS, 506, 3024
Henry A. et al., 2021, ApJ, 919, 143
Jafariyazani M., Newman A. B., Mobasher B., Belli S., Ellis R. S., Patel S. G., 2020, ApJ, 897, L42
Johnson J. W., Weinberg D. H., 2020, MNRAS, 498, 1364
Kriek M. et al., 2016, Nature, 540, 248
Kriek M. et al., 2019, ApJ, 880, L31
Larson R. B., 1974, MNRAS, 169, 229
Lonoce I., Maraston C., Thomas D., Longhetti M., Parikh T., Guarnieri P., Comparat J., 2020, MNRAS, 492, 326
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017, MNRAS, 466, 4780
McDermid R. M. et al., 2015, MNRAS, 448, 3484
Maiolino R., Mannucci F., 2019, A&AR, 27, 3
Marinacci F. et al., 2018, MNRAS, 480, 5113
Matteucci F., 2012, Chemical Evolution of Galaxies. Springer, Berlin
Matthee J., Schaye J., 2018, MNRAS, 479, L34
Mobasher B. et al., 2015, ApJ, 808, 101
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, Comput. Astrophys. Cosm., 6, 2
Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192
Pillepich A. et al., 2018a, MNRAS, 473, 4077
Pillepich A. et al., 2018b, MNRAS, 475, 648
Prohaska T. et al., 2022, Pure Appl. Chem., 94, 573
Springel V. et al., 2018, MNRAS, 475, 676
Taylor P., Kobayashi C., 2016, MNRAS, 463, 2465
Terrazas B. A. et al., 2020, MNRAS, 493, 1888
Thomas D., Maraston C., Bender R., 2003, MNRAS, 339, 897
Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775
Tinsley B. M., 1980, Fund. Cosmic Phys., 5, 287
Torrey P. et al., 2019, MNRAS, 484, 5587
Trager S. C., Faber S. M., Worthey G., González J. J., 2000, AJ, 120, 165
Trayford J. W., Schaye J., 2019, MNRAS, 485, 5715
Tremonti C. A. et al., 2004, ApJ, 613, 898
van Dokkum P., Conroy C., Villaume A., Brodie J., Romanowsky A. J., 2017, ApJ, 841, 68
van der Wel A. et al., 2016, ApJS, 223, 29
van der Wel A. et al., 2021, ApJS, 256, 44
van Loon M. L., Mitchell P. D., Schaye J., 2021, MNRAS, 504, 4817
Vogelsberger M. et al., 2014a, MNRAS, 444, 1518
Vogelsberger M. et al., 2014b, Nature, 509, 177

Weinberg D. H. et al., 2019, ApJ, 874, 102
Weinberger R. et al., 2017, MNRAS, 465, 3291
Zibetti S., Gallazzi A. R., 2022, MNRAS, 512, 1415

APPENDIX A: STELLAR ABUNDANCES AT $z = 0.7$

In this appendix, we explore the stellar abundances of TNG galaxies at $z = 0.7$. Fig. A1 shows the [Mg/H], [Fe/H], and [Mg/Fe] abundances as a function of stellar mass (left-hand panel) and effective size (right-hand panel). The trends are qualitatively similar to those shown in Fig. 2 for the $z = 2$ population, except for a negative slope in the relation between [Fe/H] and stellar mass (at $z = 2$ this relation is approximately flat, although its slope is formally negative as well). As discussed in Section 3.3, the slope of these relations becomes progressively shallower with cosmic time, and the scatter increases. In the figure, we also show the measured abundances for observed galaxies. The black marker is from a stack of dozens of galaxies at $z \sim 1.15$ observed by Carnall et al. (2022), and does not have a meaningful measurement for the effective size. The purple points are individual galaxies observed at $0.6 < z < 0.75$ by the LEGA-C survey (van der Wel et al. 2016), for which Beverage et al. (2021) measured stellar abundances. The higher quality of spectroscopic data that can be obtained for galaxies at $z = 0.7$ compared to observations at higher redshifts enables a more detailed comparison between simulations and observations. Generally, we find that measured abundances appear to be 0.2–0.3 dex lower than the simulated values. However, such direct comparison does not take into account several biases, as

discussed in Section 5.4; moreover, the most recent LEGA-C data release (van der Wel et al. 2021) revealed a potential issue with the background subtraction in earlier data releases that may cause the published measured abundances to be systematically underestimated by 0.1–0.2 dex (Beverage, private communication). What is most striking in Fig. A1 is the fact that simulations and observations follow very similar trends. Most notably, both [Mg/H] and [Fe/H] decline with R_e in the Beverage et al. (2021) sample, with a slope that is fully consistent with the TNG result. This finding supports the existence of a nearly universal abundance profile in real galaxies, which would naturally explain the observed trends.

The same observations are compared to the TNG results on the [Mg/Fe] versus [Fe/H] plane in Fig. A2. For an easy comparison, we include the quiescent $z = 2$ trend line in each panel. Here, too, an overall offset towards lower abundances is present in the observations. The scatter is large, and no sequence is evident in the data. The Beverage et al. (2021) galaxies have also been colour-coded according to their measured stellar ages (first panel) and effective sizes (last panel); no measurements of the formation time-scale are available. The measured stellar ages do not follow the same trend apparent in the TNG simulations; the size trend, on the other hand, is consistent with the TNG prediction, where larger galaxies tend to have lower [Fe/H] but higher [Mg/Fe]. We caution, however, that subtle biases in the observations due to the requirement of high signal-to-noise ratio on the spectra, which generally favors compact and young systems, may affect the comparison with the simulations.
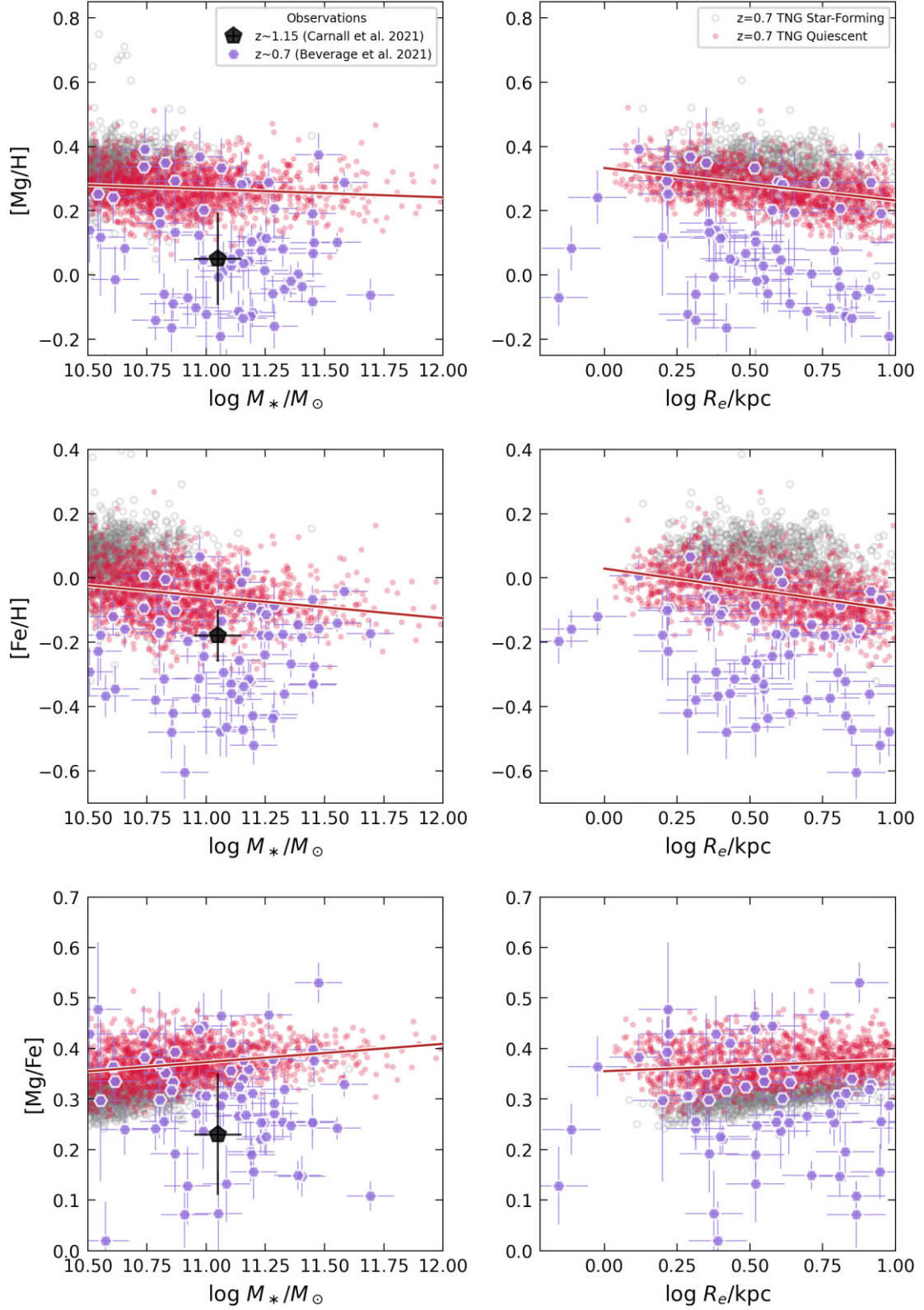
Figure A1. Stellar abundance ratios for galaxies at intermediate redshift as a function of mass (left-hand panel) and effective radius (right-hand panel). TNG galaxies at $z = 0.7$ are shown as small circles, with quiescent systems in red and star-forming galaxies in grey. In each panel, the red line is a linear fit to the simulated quiescent galaxies. Larger symbols mark observed galaxies (Beverage et al. 2021; Carnall et al. 2022).
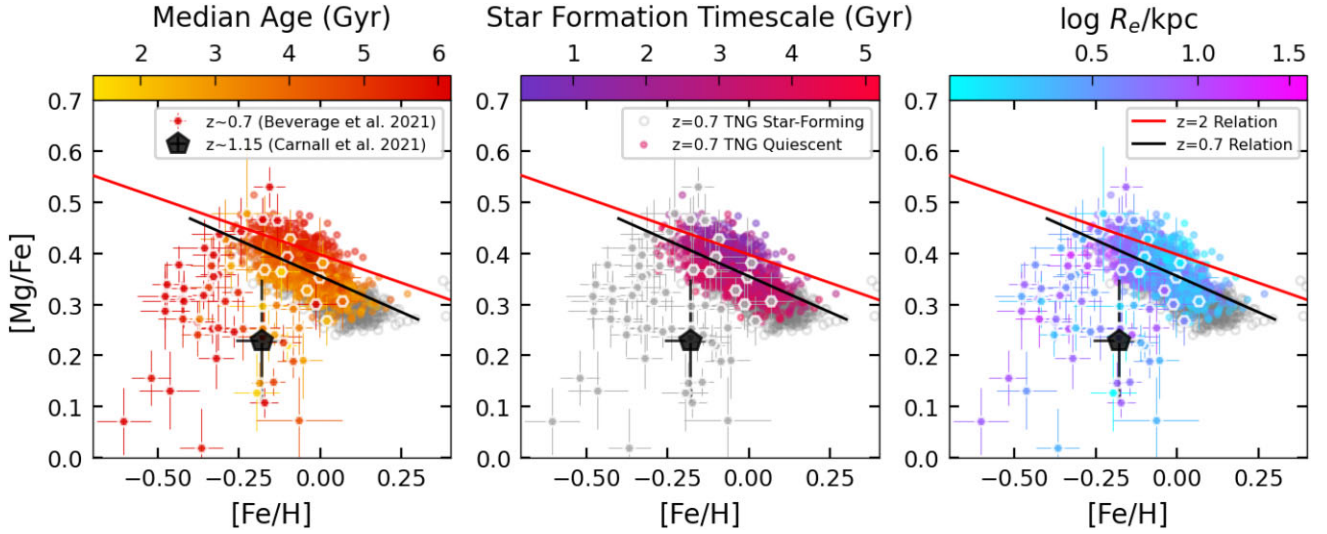
Figure A2. [Mg/Fe] versus [Fe/H] for $z = 0.7$ simulated galaxies. In each panel, quiescent galaxies are colour-coded according to a different variable: the median stellar age, the star-formation time-scale (both calculated within one effective size), and the effective size. We also show the observations at $z \sim 0.7$ from Beverage et al. (2021) (colour-coded according to the measured stellar age and effective size) and Carnall et al. (2022). The red line is a linear fit to the quiescent simulated galaxies at $z = 2$, shown in Fig. 3. The black line is a linear fit to the simulated $z = 0.7$ quiescent galaxies.

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