METALLICITY GRADIENTS IN MODERN COSMOLOGICAL SIMULATIONS: THE ROLE OF BURSTY VERSUS SMOOTH FEEDBACK AT HIGH-REDSHIFT

Alex M. Garcia[1], Paul Torrey[1]
[1]Department of Astronomy, University of Virginia, 530 McCormick Road, Charlottesville, VA 22904

12 March 2026

ABSTRACT

The distribution of gas-phase metals within galaxies encodes the impact of stellar feedback on galactic evolution. At high-redshift, when galaxies are rapidly assembling, feedback-driven outflows and turbulence can strongly reshape radial metallicity gradients. In this work, we use the FIRE-2, SPICE, Thesan and Thesan Zoom cosmological simulations – spanning a range of stellar feedback from bursty (time-variable) to smooth (steady) – to investigate how these feedback modes shape gas-phase metallicity gradients at $3 < z \lesssim 11$. Across all models, we find that galaxies with bursty feedback (FIRE-2, SPICE Bursty, and Thesan Zoom) develop systematically flatter (factors of $\sim 2-10$) metallicity gradients than those with smooth feedback (SPICE Smooth and Thesan Box). These results demonstrate that bursty stellar feedback provides sufficient turbulence to prevent strong negative gradients from forming, while smooth stellar feedback does not generically allow for efficient radial redistribution of metals thereby keeping gradients steep. Finally, we compare with recent observations, finding that the majority – but, notably, not all – of the observed gradients may favor a bursty stellar feedback scenario. In all, these results highlight the utility of high-resolution observations of gas-phase metallicity at high-redshift as a key discriminator of these qualitatively different feedback types.

1 INTRODUCTION

Feedback is central to galaxy evolution, but capturing its impact on galaxies remains a major challenge for simulations. Stars are among the most important sources of feedback, driving it through multiple channels including supernovae, photoionization and radiation pressure, protostellar jets, and stellar winds (e.g., Jijina & Adams 1996; Yorke et al. 1989; Dale et al. 2005; Quillen et al. 2005; Evans et al. 2009; Lopez et al. 2011). These processes can influence properties on the galactic scales, yet modern galaxy evolution simulations lack the resolution to directly model the micro-physics of stellar feedback and turbulence within the interstellar medium (ISM; Crain & van de Voort 2023). Nonetheless, accurately modeling feedback on these unresolved scales is imperative to producing realistic galaxy populations.

At present, galaxy evolutionary simulations that model baryonic feedback can be broadly categorized as either: *(i)* high-resolution zoom-in simulations of individual galaxies and their local environment (~a few Mpc scales) and *(ii)* lower-resolution large ($\gtrsim 35-300$ Mpc scales) cosmological box simulations (see Vogelsberger et al. 2020, for a review).

High-resolution simulations (e.g. FIRE; Hopkins et al. 2014, or SMUGGLE; Marinacci et al. 2019) can begin to resolve the multi-phase ISM down to approximately giant molecular cloud (10^4 $M_\odot$) scales, allowing for explicit modeling of star formation, feedback, and cooling. Such simulations have been successful at reproducing a number of galactic-scale properties (El-Badry et al. 2016; Sparre et al. 2017; Ma et al. 2017; Torrey et al. 2017, etc.). However, owing to the high computational cost, it is difficult to produce large samples of galaxies across a variety of cosmological environments (although, see recent efforts by FIREbox, Feldmann et al. 2023, and COLIBRE, Schaye et al. 2025). Lower-resolution large-box simulations, on the other hand, provide a large population of galaxies in a wide diversity of environments (Vogelsberger et al. 2014; Schaye et al. 2015; Davé et al. 2019; Pillepich et al. 2018, etc.). The cosmological box simulations have had their share of success reproducing population-level scaling relationships (e.g., Agertz et al. 2011; Genel et al. 2014; Sparre et al. 2015; Ferrero et al. 2017; Davé et al. 2019; Torrey et al. 2019), but struggle at high-redshift and small scales (Sun et al. 2023; Shen et al. 2024; Qi et al. 2025). The fundamental limitation of these models is that the scales of star formation and feedback are entirely unresolved. These processes are therefore usually treated entirely as phenomenological prescriptions. Typical treatments rely on simplifying assumptions, such as describing the behavior of the dense, star-forming ISM with an effective equation of state (e.g., Springel & Hernquist 2003; Schaye & Dalla Vecchia 2008) or temporarily decoupling feedback-driven winds from the hydrodynamics to help facilitate the escape of gas (Vogelsberger et al. 2013; Davé et al. 2019).

A key consequence of different treatments of star formation and feedback is that the star formation histories of galaxies in explicit models are much more bursty (strongly time variable) than those of equation of state ISM models. The bursty nature of the feedback from explicit models comes from their self-regulation of the star-forming ISM. While an equation of state provides additional pressure support that prevents the collapse of all gas into stars (Springel & Hernquist 2003; Schaye & Dalla Vecchia 2008; Ploeckinger et al. 2024, 2025; Burger et al. 2025), such pressure support is not included out-of-the-box in explicit ISM models. Therefore, the treatment of feedback needs to prevent all of the local gas from converting into stars. This explicit self-regulation by feedback from stars leads to an episodic nature of star formation followed by rapid quenching.

In Garcia et al. (2025b), we show that modern large volume cosmological simulations are in tension with high-redshift ($z > 3$) observations of metallicity gradients (Troncoso et al. 2014; Wang et al.

2022; Arribas et al. 2024; Tripodi et al. 2024; Vallini et al. 2024; Venturi et al. 2024). We find that EAGLE (Schaye et al. 2015), Illustris (Vogelsberger et al. 2014), IllustrisTNG (Pillepich et al. 2018), and SIMBA (Davé et al. 2019)—notably all effective equation of state ISM models—produce systematically stronger gradients (i.e., more negative) than are observed at $z > 3$, especially in intermediate-to-high mass galaxies (10^9 M$_\odot$ < $M_\star$ < 10^{10} M$_\odot$), suggesting that these models do not mix their gas content sufficiently. Since that work, there have been a few more gradients measured at high-redshift with *JWST* at $z \gtrsim 3$ (Acharyya et al. 2025; Ivey et al. 2025; Li et al. 2025). While Li et al. (2025) presents some evidence for strong negative gradients at high-redshift, the preference for flat gradients at high-redshift persists. If galaxies truly do have systematically flat gradients at high-redshift, then many modern cosmological simulations do not accurately model galaxy formation in its earliest stages.

2 METHODS

We employ data products from the FIRE, SPICE, and Thesan cosmological simulations. The advantage of these models for this work is that they offer an assortment of physical implementations, both bursty and smooth stellar feedback. Moreover, each of these models includes simulations of massive galaxies ($M_\star \gtrsim 10^8$ M$_\odot$) in the early universe ($z \gtrsim 5$), allowing us to probe the earliest onset of galactic evolution. Critically, the models of this work can be broadly categorized as either having bursty or smooth stellar feedback. FIRE, SPICE Bursty, and Thesan Zoom have bursty stellar feedback, while SPICE Smooth and Thesan Box have smooth stellar feedback.

2.1 Simulations

2.1.1 FIRE

The FIRE cosmological zoom-in simulations of galaxy formation analyzed here run using the multi-physics simulation code GIZMO (Hopkins 2015) and the meshless finite mass scheme for hydrodynamics. The FIRE-2 physics model we employ here (Hopkins et al. 2018) builds upon the original FIRE model (Hopkins et al. 2014).

Star formation in the FIRE model occurs in dense, self-gravitating, Jeans-unstable molecular gas with $n_H > 1000$ cm^{-3}. FIRE assumes star formation occurs at an efficiency of 100% per free fall time; however, the star formation is regulated by feedback processes. Newly formed star particles inherit their mass and metallicity from the gas in which they formed. The stellar evolution models in FIRE come from STARBURST-99 v7.0 (Leitherer et al. 1999, 2014) assuming a Kroupa (2002) IMF with masses in the range 0.1 − 100M$_\odot$. Feedback from stars is implemented through four different channels: *(i)* radiation pressure, *(ii)* supernovae with rates set by STARBURST-99 for Type II and Mannucci et al. (2006) for Type Ia, *(iii)* stellar winds from OB and AGB stars with rates set by STARBURST-99, and *(iv)* photoelectric heating. The feedback events return mass and metals to the ISM as well as drive galactic winds and outflows. The FIRE model explicitly tracks the evolution of 11 chemical species (H, He, C, N, O, Ne, Mg, Si, S, Ca, and Fe). The yields of these species from stellar feedback events come from van den Hoek & Groenewegen (1997), Marigo (2001), and Izzard et al. (2004) for stellar winds, Nomoto et al. (2006) for core-collapse supernovae, and Iwamoto et al. (1999) for Type Ia supernovae. All of the FIRE simulations analyzed in this work implement an explicit model for unresolved turbulent metal diffusion (Hopkins 2017; Su et al. 2017; Escala et al. 2018).

We use data products from the FIRE-2 public data release (Wetzel et al. 2023, 2025). For the analysis of FIRE, we use both the GIZMO ANALYSIS and HALO ANALYSIS tools provided by Wetzel et al. (2016).

2.1.2 SPICE

We also use data products from the SPICE cosmological simulations (Bhagwat et al. 2024). SPICE is a series of cosmological box simulations run with the RAMSES-RT code (Rosdahl et al. 2013; Rosdahl & Teyssier 2015), which is an extension of the original RAMSES adaptive mesh refinement code (Teyssier 2002). The main advantages of the SPICE simulations are their variations in the treatment of stellar feedback. SPICE implements several different feedback models, two of which we investigate here: smooth-sn and bursty-sn (hereafter "SPICE Smooth" and "SPICE bursty", respectively). The SPICE Smooth model allows for stars within the particles to explode anywhere from 3–40 Myr after birth (depending on their mass and rates set by STARBURST-99), leading to smoother injection of energy back into the ISM. The SPICE Bursty model, on the other hand, assumes that all supernovae go off in a single event 10 Myr after birth. Populations of stars exploding concurrently make stellar feedback particularly efficient in the SPICE Bursty model. It should be noted that the SPICE approach of modifying delay time distributions of supernovae is different than the bursty feedback in FIRE and Thesan Zoom. However, the key advantage of SPICE for the purposes of this work is the same physical model with different feedback implementations. Each of the SPICE boxes we analyze here have a volume of $(10\ \mathrm{cMpc}/h)^3$ at initial baryon mass resolution of 975 M$_\odot$.

2.1.3 Thesan

Finally, we use data products from the Thesan suite of simulations. Thesan consists of a large $(95.5\ \mathrm{cMpc})^3$ box (Garaldi et al. 2022; Kannan et al. 2022; Smith et al. 2022, hereafter Thesan Box), as well as a series of high resolution zoom-in simulations (Kannan et al. 2025, hereafter Thesan Zoom). Both the box and zooms are run using the radiation hydrodynamics moving-mesh code AREPO-RT (Kannan et al. 2019; based on the original AREPO code Springel 2010). Both Thesan Box and Thesan Zoom add an additional model for the creation and destruction of dust from gas according to McKinnon et al. (2016, 2017). While the details of the dust module are beyond the scope of this work, it is important to note that the dust forms out of the metals in the gas phase. Therefore, a (small) component of the metals are lost to dust formation, which we take to be negligible here.

Thesan Box

Thesan box is run using a modified version of the IllustrisTNG galaxy formation model (Pillepich et al. 2018), which incorporates the physics of both stellar and AGN feedback. Stars form in the dense ($n_H \gtrsim 0.1$ cm^{-3}) ISM according to the Springel & Hernquist (2003) equation of state and a Chabrier (2003) IMF. The Springel & Hernquist (2003) equation of state provides pressure support for the ISM and sets the efficiency at which stars form from the dense gas following an empirical Kennicutt-Schmidt relation (Schmidt 1959; Kennicutt 1998). Feedback from these stars is implemented in the form of supernovae explosions and stellar winds in the form of kinetic and thermal energy (see Pillepich et al. 2018). Black holes with mass 1.18×10^6 M$_\odot$ are seeded in halos exceeding 7.38×10^{10} M$_\odot$ in the Thesan Box model (Weinberger et al. 2018). Feedback from black holes is modeled via two channels based on the accretion rates of the

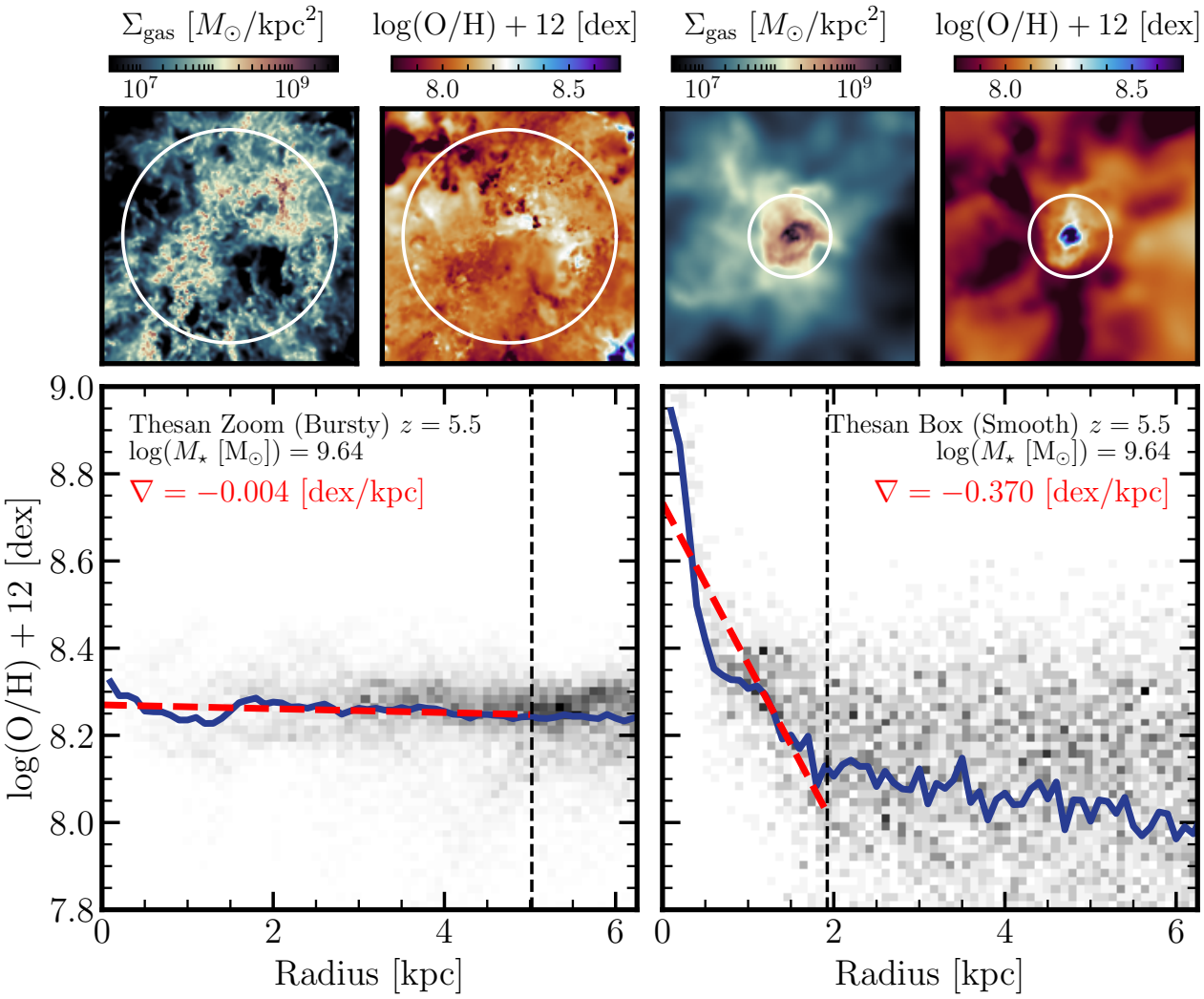

Figure 1. Example Bursty and Smooth Feedback Gradients. The gas mass surface density (top row, first and third panels) and gas-phase metallicity maps (top row second and fourth panels) within 6 kpc for a galaxy with stellar mass $10^{9.64}$ $M_\odot$ at $z = 5.5$ in Thesan Zoom (left panels) and Thesan Box (right panels) simulations. The radial distribution of metals is shown with a 2D histogram in the background of the bottom two panels. The solid blue line represents the median metallicity profile in bins of 0.1 kpc. The dashed red line shows a linear regression through the median profile at radii smaller than $R_{\rm out}$ (the radius enclosing 90% of all star formation in the galaxy). The slope of the dashed line is our reported metallicity gradient. We quote the gradient for both galaxies in the bottom panels, finding that the Thesan Zoom galaxy has a significantly flatter gradient than that of Thesan Box.

black holes: the high accretion thermal mode and low accretion kinetic mode (see Weinberger et al. 2018). The thermal mode continuously dumps thermal energy into the ISM and dominates for low mass systems, whereas high mass galaxies are dominated by the directed and pulsed kinetic winds. We utilize the highest resolution Thesan-1 volume which is a $(95.5\ {\rm cMpc})^3$ box with 2×2100^3 particles, corresponding to an initial baryon mass resolution of 3.12×10^6 $M_\odot$.

Thesan Zoom

Thesan Zoom employs a tailored version of the SMUGGLE galaxy formation model (Marinacci et al. 2019), which shares a number of similarities with FIRE. Stars form in dense ($n_{\rm H} > 10\ {\rm cm}^{-3}$), self-gravitating, Jeans-unstable gas with an efficiency per free-fall time of 100%. Similar to FIRE, the global star formation is self-regulated by feedback from the newly formed stars and maintains a lower efficiency compared to the prescribed local value (Shen et al. 2025a; Wang et al. 2025). Stellar feedback in Thesan Zoom is modeled through a number of different channels: radiative feedback (Kannan et al. 2020), stellar winds and supernovae (Marinacci et al. 2019), and "early" stellar feedback in the form of momentum driven winds (Kannan et al. 2025). Notably, the Thesan Zoom suite includes several variations in these feedback prescriptions; however, for this work, we use only the fiducial model. We note that Thesan Zoom does not include contributions from AGN. The Thesan Zoom Suite consists of 14 halos selected from the Thesan Box run down to $z \sim 3$ at various baryon mass resolutions ranging from 142 $M_\odot$ to 9.09×10^3 $M_\odot$ (see Kannan et al. 2025 their Table 2). We make use of the highest resolution versions of each halo available (see Kannan et al. 2025 their Table 3).

2.2 Galaxy Selection Criteria

The simulations tend to have different resolutions, even within the same galaxy formation model (e.g., Thesan Zoom and FIRE). To ensure that our galaxies are well-resolved, we require a threshold of $> 10^3$ gas and star particles within the system (following from Garcia et al. 2025b). Roughly speaking, this corresponds to minimum stellar and gas masses of $\sim 10^6$ $M_\odot$ in SPICE as well as high resolution FIRE and Thesan Zoom runs, $\sim 10^7$ $M_\odot$ in lower resolution FIRE and Thesan Zoom runs, and $\sim 10^8$ $M_\odot$ in the Thesan Box. Moreover, since emission line metallicity diagnostics come from star-forming HII regions (e.g., Kewley et al. 2019), we require that our galaxies have instantaneous non-zero star formation rates. Finally, we note that we restrict our analysis to the most massive subhalo within each group (i.e., the central galaxy).

2.3 Gradient Definition

We largely follow the methodology of Garcia et al. (2025b; which itself derives heavily from a combination of Ma et al. 2017 and Hemler et al. 2021) to define the metallicity gradients with only a few minor changes. We first briefly describe the methodology employed in this work, then note any changes from the previous methodology, and briefly discuss how they may impact our results. We demonstrate the application of these methodologies on two galaxies in Figure 1.

We first center the galaxy on the location of the potential minimum and rotate to a face-on orientation. To rotate the galaxy to face-on, we define the angular momentum vector of star-forming gas within a region $R_{\rm in} < r < R_{\rm out}$ (where $R_{\rm in}$ and $R_{\rm out}$ are the regions containing 5% and 90% of the total star formation of the galaxy within 10 kpc, respectively). We then align this angular momentum vector along the $+z$ axis. We note that the concept of "face-on" for galaxies with bursty feedback is not as well posed as for the more disk-like galaxies of smooth feedback models (see, e.g., Figure 1). To make as fair a comparison as possible, we re-orient all galaxies in each model according to the above prescription. Since the ISM of bursty feedback galaxies lacks coherent structure, we do not anticipate that the (re-)orientation plays a significant role in our results.

With our face-on galaxy, we create two-dimensional mass-weighted metallicity maps and gas mass maps using pixel sizes of 0.1 kpc × 0.1 kpc (shown in the top row of Figure 1). We then convert the two-dimensional maps into one-dimensional radial distributions (gray background distribution in large panels of Figure 1). We further reduce the radial distribution of metals into a singular median metallicity profile in bins of 0.1 kpc (solid blue lines in large panels of Figure 1). Finally, we summarize the median metallicity profile by fitting a linear regression (in logarithmic metallicity) in the region $r < R_{\rm out}$ (dashed red lines in Figure 1).

3 RESULTS

3.1 Distribution of Metallicity Gradients in each simulation

We begin by characterizing the metallicity gradients in each simulation. Figure 2 shows the probability density function (PDF) of all of the measured gradients in the bursty feedback models (top left is Thesan Zoom, top middle is SPICE Bursty, and top right is FIRE) and the smooth feedback models (bottom left is Thesan Box and bottom middle is SPICE Smooth). We note that combining the gradients in this way is a crude comparison, since it ignores several important factors, e.g., the redshift of the galaxy or its stellar mass. Despite the crudeness of the comparison, we find a qualitative difference in the gradients measured in the smooth feedback models to those of the bursty feedback models, with the latter typically having flatter gradients.

To quantify the discrepancy between the two different types of simulations, the bottom right panel of Figure 2 shows the median and spread (taken to be the 16$^{\rm th}$ and 84$^{\rm th}$ percentiles of the distributions) for each model analysed in this work. We find that the median gradients are $-0.021^{0.008}_{-0.072}$ dex/kpc in Thesan Zoom, $-0.124^{0.048}_{-0.333}$ dex/kpc in SPICE Bursty, $-0.016^{0.022}_{-0.066}$ dex/kpc in FIRE, $-0.205^{-0.124}_{-0.373}$ dex/kpc in Thesan Box, and $-0.248^{-0.037}_{-0.553}$ dex/kpc in SPICE Smooth. On the whole, the bursty feedback models produce significantly flatter than their smooth counterparts: for the Thesan simulations, the gradients in the Box are approximately an order of magnitude steeper than their Zoom counterparts, while the SPICE Smooth gradients are steeper than the SPICE Bursty by a factor of ~ 2. Moreover, the width of the distribution of gradients in the bursty models is significantly smaller than those of the smooth models, with the exception of SPICE which have comparable scatter. In each of the samples, we find that the distributions tend to be skewed towards more negative gradients, with the 16$^{\rm th}$ percentiles being further from the median than the 84$^{\rm th}$ percentiles. Clearly, even without more careful considerations for the detailed properties of the individual systems, there is a qualitative difference between the metallicity gradients from bursty and smooth feedback models. The bursts of feedback from stars seem to efficient redistributing metals through the galaxy, whereas the smooth feedback models do not.

It is also noteworthy that the gradients are remarkably consistent across similar models, regardless of the details of the simulation implementation. This similarity echoes the low redshift, large box simulations we investigated previously (Garcia et al. 2025b). While each simulation model—both in this work and in Garcia et al. (2025b)—is attempting to model the same general physics (e.g., star formation, stellar feedback, chemical enrichment), there is a wide diversity of numerical implementations and assumptions that go into the different models (see, e.g., Section 2.1). Yet, despite these differences, the metallicity gradients in each of the models have highly self-similar trends: smooth feedback models have strong negative gradients while bursty feedback models have flatter gradients (this similarity is also shared with the lower redshift simulation volumes, which we discuss in more detail in Section 4.1). There is, of course, variation between the predictions when considering properties of the host galaxies (which we discuss more in the next section); however, the level of out-of-the-box agreement is worth appreciating. The agreement is especially remarkable given that the total metal budget of galaxies in different simulation models can vary substantially (Garcia et al. 2024b, 2025a). Thus, we emphasize the utility of metallicity gradients as a diagnostic for stellar feedback at high-redshift ($z \gtrsim 3$), which we discuss in more detail in Section 4.1.

4 DISCUSSION

4.1 13 Billion Years of Gradient Evolution from Cosmological Simulations

Figure 3 shows the redshift evolution of metallicity gradients in nine modern cosmological simulation models. We categorize each simulation model as having either bursty or smooth stellar feedback. The bursty feedback models are FIRE-2 (Hopkins et al. 2018), Thesan Zoom (Kannan et al. 2025), and SPICE Bursty (Bhagwat et al. 2024), while the smooth feedback models are EAGLE (Schaye et al. 2015), Illustris (Vogelsberger et al. 2014), IllustrisTNG (Pillepich

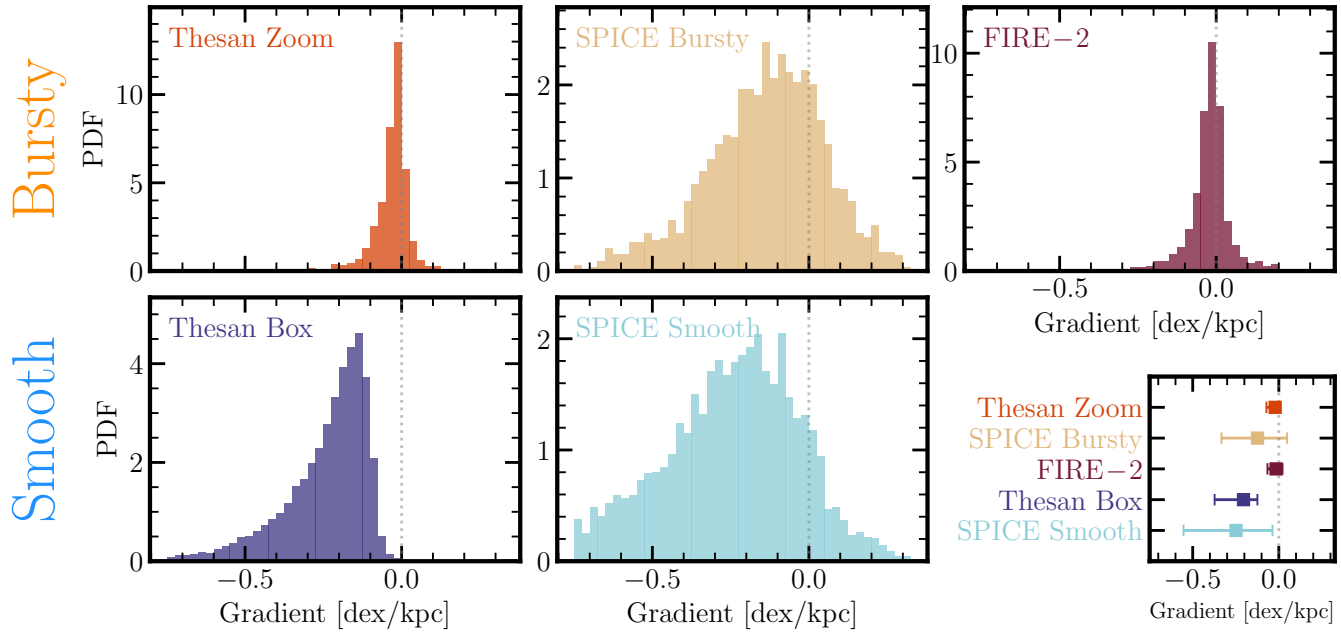

Figure 2. Distribution of Gas-Phase Radial Metallicity Gradients in Bursty and Smooth Feedback Models at High-Redshift ($3 < z \lesssim 11$)**.** Our sample of metallicity gradients in Thesan Zoom (top left), SPICE Bursty (top middle), FIRE-2 (top right), Thesan Box (bottom left), and SPICE Smooth (bottom middle). The top row represent the bursty feedback models, while the bottom row represents the smooth feedback models. We summarize the distributions in the small figure in the bottom right with the median, 16$^{\rm th}$ percentile, and 84$^{\rm th}$ percentile of the distributions. We find that the gradients in the smooth feedback models are systematically more negative than their bursty feedback counterparts.

et al. 2018), SIMBA (Davé et al. 2019), Thesan Box (Garaldi et al. 2024), and SPICE Smooth (Bhagwat et al. 2024). The data from the EAGLE, Illustris, TNG, and SIMBA simulations come from Garcia et al. (2025b; see caveat about slight change in methodology in Section 2.3). We note that this plot also includes FIRE-2 galaxies from the "core suite". We treat these systems the same as their high-redshift counterparts and only report gradients at integer and half-integer redshifts at $z < 5$. Briefly, we find that our analysis of the FIRE-2 core suite at $z < 5$ is consistent with previous FIRE papers, finding a diversity of metallicity gradients with a preference for flat gradients (Ma et al. 2017; Bellardini et al. 2021; Graf et al. 2024; Sun et al. 2024a).

Figure 3 is essentially an updated version of Gibson et al. (2013; their Figure 1), which included the evolution of four galaxies to $z \sim 2.5$. Gibson et al. (2013) find that the systems with 'enhanced' feedback have flat gradients, while systems with 'conservative' feedback have strong negative gradients that get stronger further back in time. Remarkably, we find qualitatively the same trend with the benefit of several additional simulation models with updated physics models and several orders of magnitude more galaxies. Galaxies with bursty feedback tend to have weaker metallicity gradients than their smooth feedback counterparts, especially at high-redshift ($z \gtrsim 2$) and with large stellar masses (Garcia et al. 2025b, Figure 3). The rate at which we find gradients to steepen with redshift in the smooth feedback models is approximately -0.02 dex/kpc/Δz (in good agreement with the values of Hemler et al. 2021; Garcia et al. 2025b), while the bursty feedback models are roughly consistent with 0.00 dex/kpc/Δz.

We note that the strength of gradients appears to "taper" at the highest redshifts ($z \gtrsim 8$) in the smooth feedback models, instead of continuing the trend of stronger gradients with increasing redshift. We argue that this is likely a selection effect of the stellar masses within the samples at each redshift. Lower mass galaxies ($M_\star \lesssim 10^9$ M$_\odot$) in smooth feedback models tend to have flatter gradients than their high-mass counterparts (although it is unclear if this holds for galaxies at all redshifts for very massive $M_\star > 10^{11}$ M$_\odot$ galaxies; see Garcia et al. 2025b). The sample of galaxies in, e.g., Thesan Box and SPICE Smooth tends towards lower mass systems at these very early times. It is worth noting that this same effect was noticed previously within the TNG suite, where the composite evolution of the gradients depended more on the mass composition of the sample than, e.g., resolution considerations (see discussion in Appendix A and Figure 5 of Garcia et al. 2025b). We therefore suggest that, given a sufficiently large volume, it is likely that the gradients would continue to steepen at redshifts $z > 8$ and that the observed tapering is just a selection effect from the cosmological volumes.

4.2 Comparison with High-Redshift ($z > 3$) Observations

We now make a detailed comparison to observed gas-phase metallicity gradients, with an emphasis on high-redshift ($z > 3$). We provide a more detailed comparison with low redshift observations in the smooth feedback models in Section 4.3.1 of Garcia et al. (2025b). We note that the FIRE-2 core suite also spans $z < 5$, which we did not discuss previously. Briefly, we find the lower redshift FIRE-2 sample to be broadly consistent with the wide diversity of gradients observed at low redshift (see also Ma et al. 2017; Sun et al. 2024a, for more detailed comparisons of metallicity gradients at $z < 5$ in FIRE-2 galaxies).

There have been several observations of metallicity gradients at redshifts $z > 3$ to date (Troncoso et al. 2014; Arribas et al. 2024; Tripodi et al. 2024; Vallini et al. 2024; Venturi et al. 2024; Rodríguez Del Pino et al. 2024; Acharyya et al. 2025; Fujimoto et al. 2025; Ivey

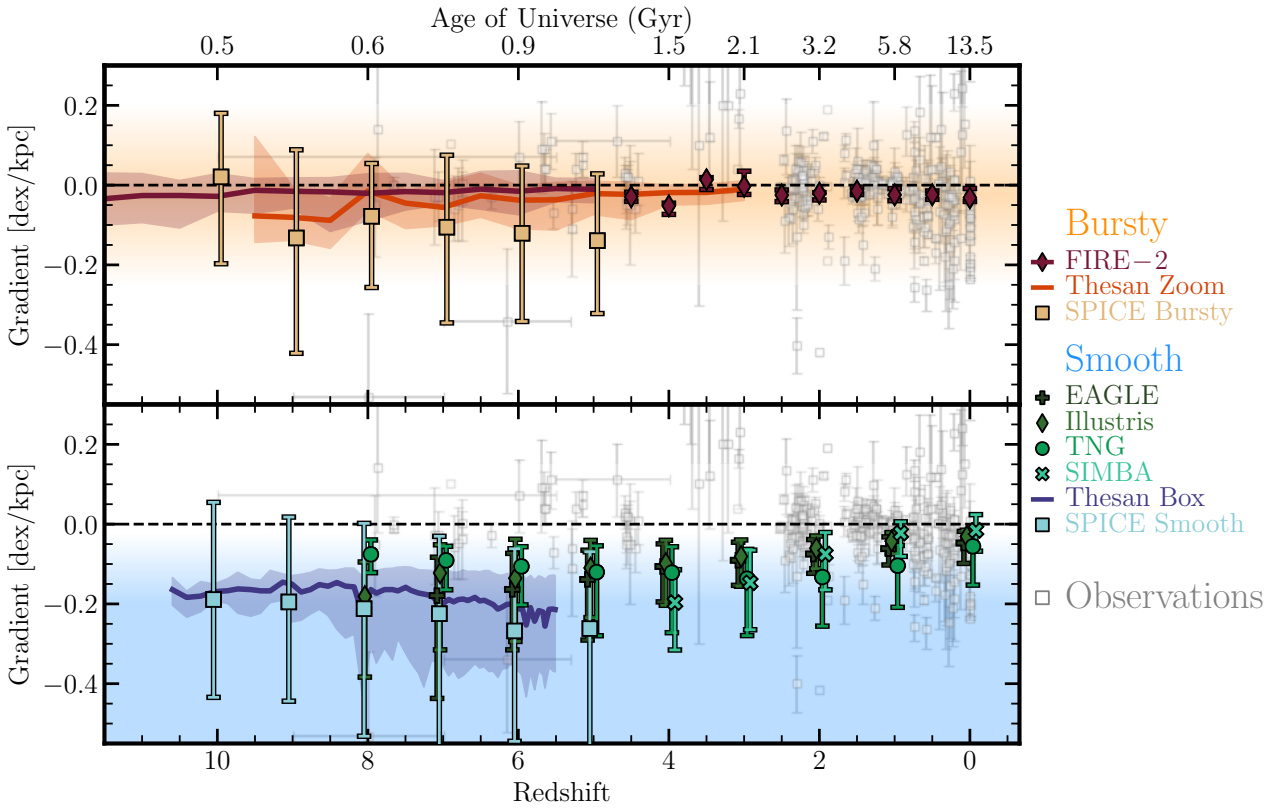

Figure 3. 13 Billion Years of Gas-Phase Metallicity Gradient Evolution in Cosmological Simulations. The composite evolution of metallicity gradients from this work and Garcia et al. (2025b) in (**Top**) bursty feedback models FIRE-2 (maroon diamonds), Thesan Zoom (solid orange line), and SPICE Bursty (orange squares), as well as (**Bottom**) smooth feedback models EAGLE (plus), Illustris (green diamonds), TNG (circles), SIMBA (Xs), Thesan Box (solid blue line), and SPICE Smooth (blue squares). The lines/markers represent the median gradient at a particular redshift while the shaded regions/error bars represent the width of the 16 – 84$^{\text{th}}$ percentiles. In general, we find that bursty feedback models tend to have flatter metallicity gradients than their smooth feedback model companions, regardless of the detailed implementation of the models. The shaded regions in each panel *qualitatively* show where the majority of each feedback type's gradients lie, particularly at high-redshift. We also present a compilation of observed metallicity gradients from Rupke et al. (2010); Queyrel et al. (2012); Swinbank et al. (2012); Jones et al. (2013, 2015); Troncoso et al. (2014); Leethochawalit et al. (2016); Wang et al. (2017, 2019, 2022); Carton et al. (2018); Förster Schreiber et al. (2018); Curti et al. (2020); Grasha et al. (2022); Li et al. (2022, 2025); Arribas et al. (2024); Tripodi et al. (2024); Vallini et al. (2024); Venturi et al. (2024); Acharyya et al. (2025); Fujimoto et al. (2025); Ivey et al. (2025); Ju et al. (2025); Valé et al. (2025). Recent high-redshift observations, with the exception of those from Li et al. (2025), seem to favor more bursty feedback scenarios.

et al. 2025; Li et al. 2025), with the majority coming from *JWST* in the last few years. We show these observational metallicity gradients in Figure 3 (noting that Figure 3 also contains observational gradients from $z < 3$; which are discussed in detail in Garcia et al. 2025b). We suspect, however, that the minor overlap in stellar mass does not significantly influence our comparison. On the whole, the current consensus at high-redshift shows a preference for flatter metallicity gradients (which can be generally seen in Figure 2); however, stacked observations from Li et al. (2025) find very strong negative gradients (which we discuss in more detail in Section 4.2.1). This picture seems to qualitatively align with the bursty feedback models presented in this work (although it does not necessarily suggest that feedback in the observed Universe is necessarily as bursty as these simulations; see discussion in Section 4.2.2).

Furthermore, it is worth acknowledging that a recent high redshift observations suggest that a significant fraction of systems (perhaps as high as 45%) show evidence of being disks (e.g., Ferreira et al. 2022; Sun et al. 2024b). The bursty feedback models tend to have a more chaotic, clumpy ISM due to the episodic blowouts of gas, whereas the smooth feedback models tend to form disk-like structures earlier (Shen et al. 2024, 2025b). Despite the qualitative disagreement between the (lack of) structure in the bursty models, the current observations seem to favor flatter gradients. While more careful, detailed comparisons are needed to measure the extent to which the structure in the different feedback models matches the observations, it is an intriguing tension that current models cannot, en masse, reproduce both structure at early times and flattened (or inverted) metallicity gradients.

4.2.1 *On Observational Systematics*

It should be noted that there are a number of observational systematics that may influence the measure of metallicity gradients of galaxies. Here, we briefly discuss some of the limiting factors and how they may systematically impact observed samples.

The choice of metallicity calibration, for one, can have a significant impact on the measured gradient of observed galaxies. As a concrete example, Ivey et al. (2025) find that the measured gradient of their galaxy ID6355 goes from -0.01 ± 0.01 dex/kpc to -0.11 ± 0.03 dex/kpc with changes from strong line to direct electron temperature diagnostics, respectively.

Another important consideration – given the compact size of the

high-redshift systems we analyze here (Shen et al. 2024; McClymont et al. 2025a) – is the impact of spatial resolution. Yuan et al. (2013) show that flatter gradients can be obtained just by downgrading the angular resolution of the observations. Worse angular resolution can create a smearing effect wherein the emission lines are weighted towards regions of stronger emission, potentially causing an overestimate of the metallicity in the outskirts of the galaxy and flattening a derived gradient. Moreover, Poetrodjojo et al. (2019) show that the contribution of diffusion ionized gas (DIG) can also significantly flatten gradients low resolution observations. Metallicity diagnostics are generally developed for resolutions on the order of single HII regions (Kewley et al. 2019), yet current high-redshift observations are likely not resolving these small scales within the ISM. The addition of DIG, which contributes more significantly further from the center of galaxies (Kewley et al. 2019), can cause commonly used metallicity diagnostics to be overestimated at large radii, flattening gradients. Careful considerations for spatial resolution are thus required to robustly measure metallicity gradients at high-redshift. To date, several works have leveraged the power of *JWST* to measure metallicity gradients (e.g., Arribas et al. 2024; Venturi et al. 2024; Acharyya et al. 2025; Fujimoto et al. 2025; Li et al. 2025), mostly finding flat gradients (with the exception of stacked gradients from Li et al. 2025). It is interesting to note that the stacked observations are the only measurements that give strong negative gradients at $z > 3$; however, it is presently unclear if stacking gradients in this way systematically biases the result towards more negative gradients (see Li et al. 2025, for a more complete discussion of their stacking procedure). Regardless, future high-resolution efforts that make considerations for the spatial resolution and contributions from DIG to obtain metallicity gradients at high-redshift could potentially help discriminate between the different feedback models.

Another consideration is that, in the smooth feedback models, the star-forming regions of galaxies at high-redshift are highly compact with strong negative gradients that flatten significantly outside this region (see Figure 1 for an example in Thesan Box; see also Garcia et al. 2023, Tapia-Contreras et al. 2025). This gradient flattening in the smooth feedback models is set by the competition between gas mixing and enrichment at large radii (Garcia et al. 2023). Observations of massive galaxies with poor spatial resolution could potentially wash out the small, high metallicity interiors of these galaxies and find a flatter gradient than we report here.

4.2.2 *On Simulation Modeling Limitations*

If the observations of flattened gradients are robust to the above systematics, there is evidence that the smooth feedback models may not sufficiently mix the metal content within their ISM. While rapid expulsions of gas driven by bursty stellar feedback is one method by which this mixing can be driven, it is not necessarily the only method. For example, diffusion of metals from unresolved turbulence within the ISM could play a significant role in redistributing metals and flattening gradients. Using the FIRE-2 model for this unresolved turbulence (described in Su et al. 2017; Escala et al. 2018; Hopkins et al. 2018), Bellardini et al. (2021) show that changes to the diffusion coefficients can systematically impact the radial distribution of metals. Specifically, those authors find that increased metal diffusion flattens metallicity gradients. It is therefore possible that the smooth feedback models require increased metal mixing from these unresolved eddies.

Increased turbulent diffusion mixes metals, but likely would not significantly change the overall total metal content. The existence—or lack thereof—of a correlation between the scatter about the mass-metallicity relation (MZR) and star formation rates (often referred to as the Fundamental Metallicity Relation; FMR) would therefore likely also be unimpeded in smooth feedback models with increased subgrid metal diffusion. While there is evidence that the FMR evolves in the smooth feedback models (i.e., the FMRs are weak and/or dynamic; Garcia et al. 2024b, 2025a), there still exists an anti-correlation between metallicity and SFR up to at least $z = 8$. Bursty feedback, on the other hand, should disrupt the interplay of gas accretion and metal return that gives rise to this anti-correlation (Garcia et al. 2024a,b, 2025a; Bassini et al. 2024; McClymont et al. 2025b). The existence of FMR-like anti-correlations between metallicity and SFR in the high-redshift Universe—regardless of strong/weak or static/dynamic designation—could provide further constraining power as to whether flattened gradients are suggestive of bursty feedback or increased small-scale turbulent eddies within the ISM.

5 CONCLUSIONS

Our key conclusions are as follows:

- We find that metallicity gradients are systematically flatter (factors of $\sim 2-10$) in bursty (FIRE-2, SPICE Bursty, and Thesan Zoom) than in smooth (SPICE Smooth and Thesan Box) feedback models (Figure 2). We note that this result is—remarkably—mostly independent of the detailed implementations of either bursty or smooth feedback.
- We contextualize our results with the first paper in this series (Garcia et al. 2025b), which shows smooth feedback models at $0 \leq z \leq 8$, and recent *JWST*/ALMA observations (Figure 3). We find good agreement between the metallicity gradients in smooth feedback models of this work (Thesan Box and SPICE Smooth) and EAGLE, Illustris, TNG, and SIMBA, suggesting that the strong negative gradients are a highly generic feature of smooth feedback models. Moreover, the agreement between FIRE, SPICE Bursty, and Thesan Zoom suggests that the flat gradients may be a generic feature of bursty feedback models.
- Finally, we find that the recent high-redshift gradient observations (Arribas et al. 2024; Vallini et al. 2024; Venturi et al. 2024; Li et al. 2025) tend to align more closely with the gradients from bursty feedback models. This is suggestive that smooth feedback models may under-mix their metal content, either through increased feedback or a lack of a metal diffusion in the models (see discussion in Section 4.2.2); however, given current observational limitations, it is unclear how robust this finding is (see discussion in Section 4.2.1).

6 ACKNOWLEDGMENTS

AMG acknowledges support from a Virginia Space Grant Consortium Graduate STEM Research Fellowship.

REFERENCES

Acharyya A., et al., 2025, arXiv e-prints, p. arXiv:2508.05335
Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391
Arribas S., et al., 2024, A&A, 688, A146
Bassini L., Feldmann R., Gensior J., Faucher-Giguère C.-A., Cenci E., Moreno J., Bernardini M., Liang L., 2024, MNRAS, 532, L14
Bellardini M. A., Wetzel A., Loebman S. R., Faucher-Giguère C.-A., Ma X., Feldmann R., 2021, MNRAS, 505, 4586
Bhagwat A., Costa T., Ciardi B., Pakmor R., Garaldi E., 2024, MNRAS, 531, 3406

Burger J. D., et al., 2025, MNRAS,
Carton D., et al., 2018, MNRAS, 478, 4293
Chabrier G., 2003, PASP, 115, 763
Crain R. A., van de Voort F., 2023, ARA&A, 61, 473
Curti M., et al., 2020, MNRAS, 492, 821
Dale J. E., Bonnell I. A., Clarke C. J., Bate M. R., 2005, MNRAS, 358, 291
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
El-Badry K., Wetzel A., Geha M., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2016, ApJ, 820, 131
Escala I., et al., 2018, MNRAS, 474, 2194
Evans Neal J. I., et al., 2009, ApJS, 181, 321
Feldmann R., et al., 2023, MNRAS, 522, 3831
Ferreira L., et al., 2022, ApJ, 938, L2
Ferrero I., et al., 2017, MNRAS, 464, 4736
Förster Schreiber N. M., et al., 2018, ApJS, 238, 21
Fujimoto S., et al., 2025, arXiv e-prints, p. arXiv:2510.16116
Garaldi E., Kannan R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, MNRAS, 512, 4909
Garaldi E., et al., 2024, MNRAS, 530, 3765
Garcia A. M., et al., 2023, MNRAS, 519, 4716
Garcia A. M., et al., 2024a, MNRAS, 529, 3342
Garcia A. M., et al., 2024b, MNRAS, 531, 1398
Garcia A. M., et al., 2025a, MNRAS, 536, 119
Garcia A. M., et al., 2025b, ApJ, 989, 147
Genel S., et al., 2014, MNRAS, 445, 175
Gibson B. K., Pilkington K., Brook C. B., Stinson G. S., Bailin J., 2013, A&A, 554, A47
Graf R. L., Wetzel A., Bailin J., Orr M. E., 2024, arXiv e-prints, p. arXiv:2410.21377
Grasha K., et al., 2022, ApJ, 929, 118
Hemler Z. S., et al., 2021, MNRAS, 506, 3024
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., 2017, MNRAS, 466, 3387
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., 2018, MNRAS, 480, 800
Ivey L. R., et al., 2025, arXiv e-prints, p. arXiv:2507.14936
Iwamoto K., Brachwitz F., Nomoto K., Kishimoto N., Umeda H., Hix W. R., Thielemann F.-K., 1999, ApJS, 125, 439
Izzard R. G., Tout C. A., Karakas A. I., Pols O. R., 2004, MNRAS, 350, 407
Jijina J., Adams F. C., 1996, ApJ, 462, 874
Jones T., Ellis R. S., Richard J., Jullo E., 2013, ApJ, 765, 48
Jones T., et al., 2015, AJ, 149, 107
Ju M., et al., 2025, ApJ, 978, L39
Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, MNRAS, 485, 117
Kannan R., Marinacci F., Simpson C. M., Glover S. C. O., Hernquist L., 2020, MNRAS, 491, 2088
Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022, MNRAS, 511, 4005
Kannan R., et al., 2025, arXiv e-prints, p. arXiv:2502.20437
Kennicutt Jr. R. C., 1998, ARA&A, 36, 189
Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511
Kroupa P., 2002, Science, 295, 82
Leethochawalit N., Jones T. A., Ellis R. S., Stark D. P., Richard J., Zitrin A., Auger M., 2016, ApJ, 820, 84
Leitherer C., et al., 1999, ApJS, 123, 3
Leitherer C., Ekström S., Meynet G., Schaerer D., Agienko K. B., Levesque E. M., 2014, ApJS, 212, 14
Li Z., et al., 2022, ApJ, 929, L8
Li Z., et al., 2025, arXiv e-prints, p. arXiv:2506.12129
Lopez L. A., Krumholz M. R., Bolatto A. D., Prochaska J. X., Ramirez-Ruiz E., 2011, ApJ, 731, 91
Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017, MNRAS, 466, 4780
Mannucci F., Della Valle M., Panagia N., 2006, MNRAS, 370, 773
Marigo P., 2001, A&A, 370, 194
Marinacci F., Sales L. V., Vogelsberger M., Torrey P., Springel V., 2019, MNRAS, 489, 4233
McClymont W., et al., 2025a, arXiv e-prints, p. arXiv:2503.04894
McClymont W., et al., 2025b, arXiv e-prints, p. arXiv:2507.08787
McKinnon R., Torrey P., Vogelsberger M., 2016, MNRAS, 457, 3775
McKinnon R., Torrey P., Vogelsberger M., Hayward C. C., Marinacci F., 2017, MNRAS, 468, 1505
Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, Nuclear Phys. A, 777, 424
Pillepich A., et al., 2018, MNRAS, 473, 4077
Ploeckinger S., Nobels F. S. J., Schaller M., Schaye J., 2024, MNRAS, 528, 2930
Ploeckinger S., Richings A. J., Schaye J., Trayford J. W., Schaller M., Chaikin E., 2025, MNRAS, 543, 891
Poetrodjojo H., D'Agostino J. J., Groves B., Kewley L., Ho I. T., Rich J., Madore B. F., Seibert M., 2019, MNRAS, 487, 79
Qi J., et al., 2025, arXiv e-prints, p. arXiv:2501.18687
Queyrel J., et al., 2012, A&A, 539, A93
Quillen A. C., Thorndike S. L., Cunningham A., Frank A., Gutermuth R. A., Blackman E. G., Pipher J. L., Ridge N., 2005, ApJ, 632, 941
Rodríguez Del Pino B., et al., 2024, A&A, 684, A187
Rosdahl J., Teyssier R., 2015, MNRAS, 449, 4380
Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, MNRAS, 436, 2188
Rupke D. S. N., Kewley L. J., Chien L. H., 2010, ApJ, 723, 1255
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J., et al., 2015, MNRAS, 446, 521
Schaye J., et al., 2025, arXiv e-prints, p. arXiv:2508.21126
Schmidt M., 1959, ApJ, 129, 243
Shen X., et al., 2024, MNRAS, 534, 1433
Shen X., et al., 2025a, arXiv e-prints, p. arXiv:2503.01949
Shen X., Zier O., Vogelsberger M., Boylan-Kolchin M., Hernquist L., Tacchella S., Naidu R. P., 2025b, arXiv e-prints, p. arXiv:2509.19427
Smith A., Kannan R., Garaldi E., Vogelsberger M., Pakmor R., Springel V., Hernquist L., 2022, MNRAS, 512, 3243
Sparre M., et al., 2015, MNRAS, 447, 3548
Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2017, MNRAS, 466, 88
Springel V., 2010, MNRAS, 401, 791
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Su K.-Y., Hopkins P. F., Hayward C. C., Faucher-Giguère C.-A., Kereš D., Ma X., Robles V. H., 2017, MNRAS, 471, 144
Sun G., Faucher-Giguère C.-A., Hayward C. C., Shen X., Wetzel A., Cochrane R. K., 2023, ApJ, 955, L35
Sun X., et al., 2024a, arXiv e-prints, p. arXiv:2409.09290
Sun W., Ho L. C., Zhuang M.-Y., Ma C., Chen C., Li R., 2024b, ApJ, 960, 104
Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., McCarthy I. G., Crain R. A., Theuns T., 2012, MNRAS, 426, 935
Tapia-Contreras B., et al., 2025, A&A, 700, A69
Teyssier R., 2002, A&A, 385, 337
Torrey P., Hopkins P. F., Faucher-Giguère C.-A., Vogelsberger M., Quataert E., Kereš D., Murray N., 2017, MNRAS, 467, 2301
Torrey P., et al., 2019, MNRAS, 484, 5587
Tripodi R., et al., 2024, A&A, 692, A184
Troncoso P., et al., 2014, A&A, 563, A58
Valé G., et al., 2025, A&A, 701, A226
Vallini L., et al., 2024, MNRAS, 527, 10
Venturi G., et al., 2024, arXiv e-prints, p. arXiv:2403.03977
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M., et al., 2014, MNRAS, 444, 1518
Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42
Wang X., et al., 2017, ApJ, 837, 89
Wang X., et al., 2019, ApJ, 882, 94
Wang X., et al., 2022, ApJ, 938, L16
Wang Z., et al., 2025, arXiv e-prints, p. arXiv:2505.05554

Weinberger R., et al., 2018, MNRAS, 479, 4056
Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23
Wetzel A., et al., 2023, ApJS, 265, 44
Wetzel A., et al., 2025, arXiv e-prints, p. arXiv:2508.06608
Yorke H. W., Tenorio-Tagle G., Bodenheimer P., Rozyczka M., 1989, A&A, 216, 207
Yuan T. T., Kewley L. J., Rich J., 2013, ApJ, 767, 106
van den Hoek L. B., Groenewegen M. A. T., 1997, A&AS, 123, 305

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