

MEGATRON: REPRODUCING THE DIVERSITY OF HIGH-REDSHIFT GALAXY SPECTRA WITH COSMOLOGICAL RADIATION HYDRODYNAMICS SIMULATIONS

HARLEY KATZ^{1,2*}, MARTIN P. REY³, CORENTIN CADIOU⁴,
OSCAR AGERTZ⁵, JEREMY BLAIZOT⁶, ALEX J. CAMERON⁷, NICHOLAS CHOUSTIKOV⁷, JULIEN DEVRIENDT⁷, ULIANA HAUKE¹,
GARETH C. JONES^{8,9}, TAYSUN KIMM¹⁰, ISAAC LASETER¹¹, SERGIO MARTIN-ALVAREZ¹², KOSEI MATSUMOTO¹³, AUTUMN
PEARCE¹, FRANCISCO RODRÍGUEZ MONTERO^{1,2}, JOKI ROSDAHL⁶, MAHSA SANATI⁷, AAYUSH SAXENA⁷, ADRIANNE SLYZ⁷,
RICHARD STISKALEK⁷, ANATOLE STORCK⁷, OSCAR VEENEMA⁷, AND WONJAE YEE¹

¹Department of Astronomy & Astrophysics, University of Chicago, 5640 S Ellis Avenue, Chicago, IL 60637, USA

²Kavli Institute for Cosmological Physics, University of Chicago, Chicago IL 60637, USA

³University of Bath, Department of Physics, Claverton Down, Bath, BA2 7AY, UK

⁴Institut d'Astrophysique de Paris, Sorbonne Université, CNRS, UMR 7095, 98 bis bd Arago, 75014 Paris, France

⁵Division of Astrophysics, Department of Physics, Lund University, Box 118, SE-221 00 Lund, Sweden

⁶Université Claude Bernard Lyon 1, CRAL UMR5574, ENS de Lyon, CNRS, Villeurbanne, F-69622, France

⁷Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, United Kingdom

⁸Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

⁹Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK

¹⁰Department of Astronomy, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

¹¹Department of Astronomy, University of Wisconsin-Madison, Madison, WI 53706, USA

¹²Kavli Institute for Particle Astrophysics & Cosmology (KIPAC), Stanford University, Stanford, CA 94305, USA and

¹³Sterrenkundig Observatorium Department of Physics and Astronomy Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium

Version October 8, 2025

ABSTRACT

We present the MEGATRON suite of cosmological radiation hydrodynamics simulations following the formation of Milky Way-mass galaxies from the earliest cosmic epochs when Population III stars form to Cosmic Noon. The suite represents the first set of cosmological simulations that couples a vast non-equilibrium thermochemistry network of primordial species, metals, and molecules to multifrequency, on-the-fly radiation transport, allowing us to directly predict the spectral properties of early galaxies. By initializing the simulations at zero metallicity, resolving haloes well below the atomic cooling threshold, reaching parsec-scale resolution, and modeling a Milky Way-mass environment, we aim to address four key science themes: 1) Star formation at cosmic dawn, 2) Galaxy formation and the interstellar medium in the epoch of reionization, 3) The circumgalactic medium towards cosmic noon, and 4) Reionization in a local volume environment and near-field cosmology. In this introductory work, we present an overview of the physical characteristics of high-redshift MEGATRON galaxies and their environment at $z > 8$. We present a library of $> 175,000$ simulated galaxy spectra and demonstrate how the diversity of galaxy spectra seen by JWST is naturally reproduced in the context of a Λ CDM cosmology. This project represents a step towards making more direct comparisons between simulations and observations and will enable future work to both optimize methods for inferring galaxy properties from observations and to elucidate the physics that governs galaxy formation in the early Universe.

Subject headings: high-redshift galaxies, ISM, galaxy formation

1. INTRODUCTION

Prior to the 2020s, few constraints existed on the properties of the high-redshift ($z > 6$) interstellar medium (ISM). The sample of spectroscopic detections was limited to a few bright sources detected with the Hubble Space Telescope (HST) that could be targeted by the Atacama Large Millimeter Array (e.g., [Pentericci et al. 2016](#); [Bradač et al. 2017](#); [Hashimoto et al. 2019](#); [Harikane et al. 2020](#); [Carniani et al. 2020](#)) at mm wavelengths, or UV line detections from other large ground-based facilities (e.g., [Stark et al. 2017](#); [Mainali et al. 2017](#)).

The launch of the James Webb Space Telescope (JWST) catalyzed a paradigm shift in our understanding of galaxy formation in the first billion years. Designed with a large collecting area (allowing it to observe faint objects), spectroscopic capabilities (needed for precision measurements of ISM properties such as density, tem-

perature, metallicity), and a low operating temperature and mirror coatings (needed to detect infrared photons that correspond to the rest-frame UV and optical in the epoch of reionization), JWST is optimal for studies of the ISM in the first galaxies ([Gardner et al. 2006, 2023](#)).

The thousands of high-quality public JWST spectra of early galaxies are beginning to reveal population-level trends of the high-redshift ISM (e.g., [Sanders et al. 2023](#); [Cameron et al. 2023c](#); [Hu et al. 2024](#); [Topping et al. 2025](#); [Shapley et al. 2025](#)). In general, the high-redshift ISM exists at higher gas density, higher excitation, and lower metallicity compared to the low-redshift Universe. While these broad trends had been expected based on observations at intermediate redshifts (e.g., [Steidel et al. 2016](#); [Strom et al. 2017](#); [Reddy et al. 2018](#); [Shapley et al. 2019](#)) and from low-redshift “analog” galaxies (e.g., [Cardamone et al. 2009](#); [Izotov, Guseva & Thuan 2011](#); [Izotov et al. 2021](#); [Mingozi et al. 2022](#)), many surprising aspects of early galaxies have also emerged.

*E-mail: harleykatz@uchicago.edu

For example, multiple high-redshift galaxy surveys conducted with JWST have reported an overabundance of bright galaxies at $z \gtrsim 9$, a finding that contrasts with earlier HST-based constraints and many theoretical models (e.g., Finkelstein et al. 2023; Harikane et al. 2024; Leung et al. 2023; Chemerynska et al. 2023, although c.f. Willott et al. 2023). Several early galaxies also exhibit chemical abundance patterns that have never been observed in the gas-phase (e.g., Cameron et al. 2023a; Senchyna et al. 2024; Isole et al. 2023; Topping et al. 2024). Some of the high-redshift spectra appear to require exotic stellar populations or a top-heavy stellar initial mass function (e.g., Cameron et al. 2023b; Katz et al. 2024a; Cullen et al. 2025). Certain gas temperature measurements may be at odds with standard ISM heating mechanisms (e.g., Katz et al. 2023c; Laseter et al. 2024). The star formation histories of massive quiescent galaxies may place their formation at very early epochs (e.g., de Graaff et al. 2024; Glazebrook et al. 2024). Finally, a new class of compact red galaxies with peculiar “v”-shaped spectra (“Little Red Dots”) has been discovered whose exact nature remains unknown (e.g., Matthee et al. 2024a).

These findings highlight the need for updated models of high-redshift galaxy formation and the ISM. This is a significant challenge, as the thermodynamic state of the ISM for which many of the observed emission lines are exponentially sensitive is a complex interplay between heating and cooling thermochemistry processes, stellar feedback in the form of supernova (SN) explosions, stellar winds, and ionizing radiation, and large-scale galactic processes such as turbulence and inflows/outflows (e.g., Field, Goldsmith & Habing 1969; McKee & Ostriker 1977; Tielens 2005). Modifications to, for example, chemical abundance patterns or modes of star formation compared to what occurs in the solar neighborhood can lead to a vastly different ISM pressure-density distribution (e.g., Bialy & Sternberg 2019; Katz et al. 2022; Kim et al. 2023).

Historically, inferring and interpreting the physical properties of the ISM from spectra relies on line ratio measurements and photoionization modeling (see Kewley, Nicholls & Sutherland 2019 for a review). While many inferred properties from line ratios are based on quantum physics with (reasonably) well measured atomic data, photoionization modeling is subject to significant uncertainties. For example, simply changing the structure of an illuminated cloud from a 1D sphere or slab geometry to a 3D turbulent medium with the same mean density can have non-negligible impacts on the observed line ratios (e.g., Gray & Scannapieco 2017; Jin, Kewley & Sutherland 2022). Likewise, the underlying stellar population (e.g. Byler et al. 2017; Xiao, Stanway & Eldridge 2018) and included physics (cosmic rays, X-ray heating, etc, e.g., Katz et al. 2023c) can affect the observed spectrum, as can galactic-scale effects (e.g., non-linear averaging over different geometries and emission regions; Cameron et al. 2023c). Finally, there is no guarantee that the ISM is in an equilibrium state (e.g. Richings et al. 2022), which is an implicit assumption of most photoionization models. However, photoionization modeling is computationally inexpensive, allowing for large parameter spaces to be explored. Such wide parameter spaces are necessary for interpreting the diversity of

spectra seen in the real Universe.

3D numerical simulations of high-redshift galaxies represent a complementary approach to interpreting JWST spectra and images. There is no shortage of numerical simulations of early galaxy formation that predict line emission, full spectra, or imaging (e.g., Ceverino, Glover & Klessen 2017; Barrow et al. 2017; Katz et al. 2019a, 2023b; Vogelsberger et al. 2020; Arata et al. 2020; Lupi et al. 2020; Lovell et al. 2021; Trebitsch et al. 2021; Pallottini et al. 2022; Hirschmann et al. 2023; Yang et al. 2023; Kannan et al. 2022; Nakazato, Yoshida & Ceverino 2023; Kannan et al. 2025; Garg et al. 2024; Schimek et al. 2024; Bhagwat et al. 2024). However, all these simulations lack the detailed physics present in photoionization models that is needed to robustly predict spectra. More specifically, most of these simulations do not perform a detailed modeling of the ISM (e.g., Vogelsberger et al. 2020; Lovell et al. 2021; Hirschmann et al. 2023) or include radiation transport. Reducing physical fidelity allows for simulations of much larger cosmological volumes that probe more massive galaxies, and crucially, the bright end of the UV luminosity function. This tradeoff is necessary to model massive galaxies in their environments. For the subset of simulations that attempt ISM modeling of lower mass objects (e.g., Barrow et al. 2017; Katz et al. 2019a, 2023b; Trebitsch et al. 2021; Pallottini et al. 2022; Kannan et al. 2025), the non-equilibrium chemistry is typically limited to only primordial species (e.g., H, He, H₂, e⁻); hence, emission from only hydrogen and helium can be directly predicted from the simulation. None of these simulations include a detailed, non-equilibrium metal or molecular chemical network, nor have the ISM models been benchmarked against local constraints (Kim et al. 2023).

Simulations without non-equilibrium metal chemistry must be post-processed with photoionization models such as CLOUDY (Ferland et al. 2017) or Monte Carlo radiation transfer (e.g., McClymont et al. 2024) in order to model their intrinsic (before dust attenuation) spectra. Such an approach has been successful in creating realistic looking galaxies that reproduce many of the observed properties of the high-redshift galaxy population (e.g., Hirschmann, Charlot & Somerville 2023; Katz et al. 2023b; Wilkins et al. 2023; Nyhagen et al. 2024). However, numerous assumptions must be made in post-processing that limit the predictive power of the simulations. Different (reasonable) sets of assumptions can lead to significantly different mock observations. For this reason, there is strong motivation for the development of more self-consistent and predictive forward models, especially in the context of dwarf galaxies where such an approach is numerically tractable.

There are now a minority of cosmological simulation codes that can model the detailed physics from photoionization models, in non-equilibrium, in a 3D galaxy formation setting (e.g., Richings, Schaye & Oppenheimer 2014; Katz 2022; Chan et al. 2025). RAMSES-RTZ is unique among these as it is connected to on-the-fly, multi-frequency radiation transport and has been extensively benchmarked to reproduce equilibrium results computed with 1D photoionization codes (Katz 2022; Katz et al. 2024c) and the properties of the Milky Way ISM at solar metallicity (Katz et al. 2022). The non-equilibrium aspect of the code allows one to drop a major

equilibrium assumption that is ubiquitous across post-processing methods and was recently shown to fail in certain regimes (Richings et al. 2022; Ploeckinger et al. 2025). Moreover, the 3D nature of RAMSES-RTZ enables it to capture the geometric effects of H II region structure on emission lines (Gray & Scannapieco 2017; Jin, Kewley & Sutherland 2022). Hence RAMSES-RTZ can predict the intrinsic spectra¹ of simulated galaxies from cosmological initial conditions, making it uniquely suited for studying the spectral evolution of early galaxies.

In this work, we introduce the MEGATRON suite of simulations, that, for the first time, employ a detailed thermochemistry network of primordial species, metals, and molecules, coupled to on-the-fly radiation transport and a state-of-the-art galaxy formation model. The simulations represent a large zoom region around an object that collapses to a Milky Way-mass galaxy at $z = 0$ (although the simulations are stopped at much higher redshift). The MEGATRON project is designed to advance our understanding in four key science areas (Figure 1), in chronological order:

1. Predicting the demographics and spectral properties of the first generation of Population III (Pop. III) stars.
2. Quantifying how observable ISM properties respond to different galaxy and stellar evolution models at low metallicity.
3. Studying how non-equilibrium chemistry, local radiation fields, and mass assembly affect the emission and absorption observables of the circumgalactic medium (CGM) at cosmic noon.
4. Deciphering the archaeological traces left by high-redshift galaxy formation physics in the local volume environment of a Milky Way-mass galaxy.

To achieve this, the MEGATRON simulations bring together several technical improvements that build upon previous numerical efforts to model early galaxy formation. In particular:

- The explicit modeling of Pop. III star formation from zero-metallicity initial conditions facilitated by non-equilibrium H_2 cooling and a high spatial resolution ($\approx pc$). This mode of star formation is common in targeted studies of minihaloes and a select few larger-scale cosmological simulations (e.g., Wise et al. 2012; O’Shea et al. 2015; Kimm et al. 2017; Brauer et al. 2025), but remains absent in most high-redshift simulations attempting to resolve the ISM, e.g., SPHINX (Rosdahl et al. 2018), FIRE (Ma et al. 2018), THESAN-ZOOM (Kannan et al. 2025).

¹ I.e. the observed spectra prior to dust attenuation. We do not sample enough frequency bins (due to memory constraints) in the code to accurately predict the attenuated spectra on-the-fly and it must therefore be done in post-processing. Note also that when the resolution is too low, Stromgren spheres become unresolved and corrections to the temperature and ionization state must be made in post-processing (see e.g. Katz et al. 2023b). However, this only occurs in very dense gas which we empirically find is close to equilibrium.

- An energetic feedback model that was calibrated on Milky Way-mass galaxies at $z = 0$, and shown to produce a realistic stellar mass-halo mass relation at $z = 0$ (Agertz et al. 2021; Rey et al. 2023).
- A non-equilibrium chemistry model for primordial species, metals, and molecules. This is not generally captured by high-redshift simulations². Non-equilibrium chemistry can be important for accurately predicting ISM emission lines depending on the density and ionizing conditions (e.g., Richings et al. 2022; Katz et al. 2022), and its importance is expected to increase in the CGM where lower gas density leads to longer equilibrium timescales³ (Gnat & Sternberg 2007; Oppenheimer & Schaye 2013; Kumar & Chen 2025).
- An ISM thermochemistry model (Katz et al. 2022) that reproduces the conditions of the local ISM. The importance of recovering the inferred gas pressure as a function of density under such conditions was recently discussed in Kim et al. (2023).
- An emphasis on mock observations (see Figure 2), where we have aimed to reproduce (with as few additional assumptions as possible) the observing modes of NIRCам and NIRSspec on JWST.
- A grid of simulations studying the response of observables to both (i) variations in fundamental sub-grid models; and (ii) controlled variations in the mass growth history of an object using the ‘genetic modification’ approach (Roth, Pontzen & Peiris 2016; Rey & Pontzen 2018).

In this paper, we introduce the suite of MEGATRON simulations. Most of the physical ingredients of the simulations are presented in the methods paper (Katz et al. 2024c) and we outline here the numerics of two simulation suites targeting high-redshift galaxy formation and the CGM towards cosmic noon. As a validation of our approach and model, we show how the MEGATRON simulations are able to reproduce the spectral diversity seen at high redshift with JWST.

2. NUMERICAL METHODS

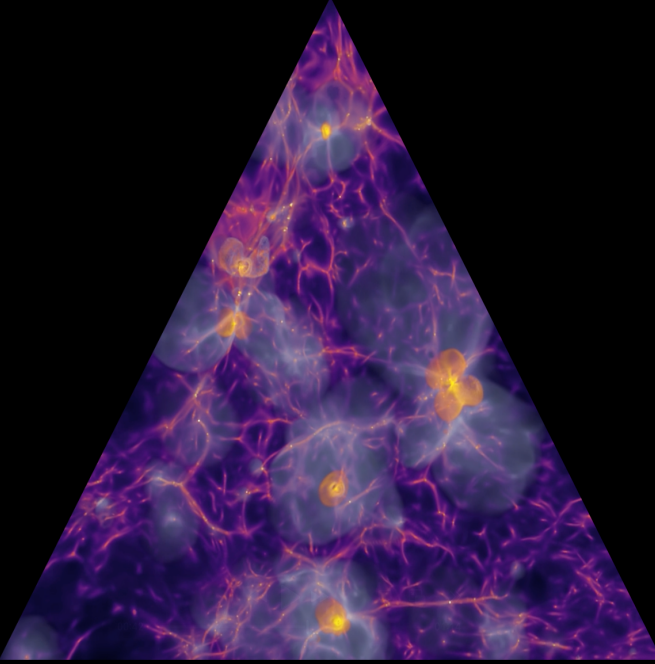
We present seven cosmological radiation hydrodynamics simulations of the Lagrange region around a Milky Way-mass galaxy. The simulations are all run with the MEGATRON galaxy formation model presented in Katz et al. (2024c) that was developed within the RAMSES-RTZ adaptive mesh refinement (AMR) code (Katz 2022). RAMSES-RTZ is a fork of RAMSES (Teyssier 2002) and RAMSES-RT (Rosdahl et al. 2013; Rosdahl & Teyssier 2015). Four simulations comprise the “high-redshift” suite, where the resolution is primarily focused on resolving the ISM in the early Universe. The physical

² Recent state-of-the-art simulations do have non-equilibrium H and He thermochemistry, but assume equilibrium for all other elements. Note also that a small minority of simulations include the low ionization states of species like carbon and oxygen (e.g., Lupi et al. 2020).

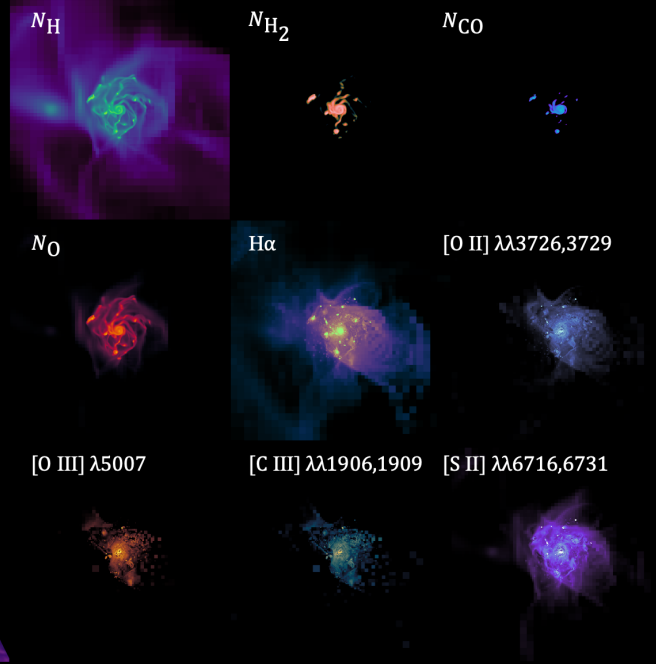
³ For example, the recombination timescale scales inversely with electron density.

Megatron: Bridging Theory and Observation in Four Dimensions

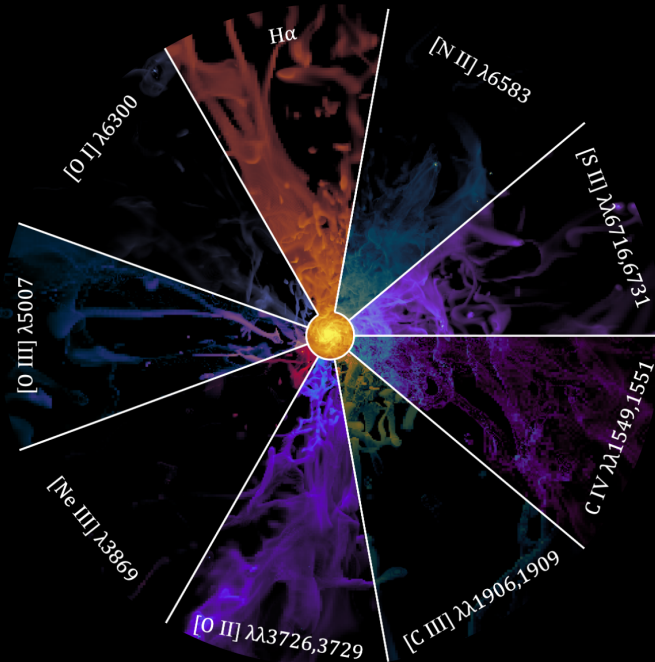
1. Star Formation at Cosmic Dawn



2. Galaxy Formation & the ISM in the EoR



3. The CGM Towards Cosmic Noon



4. Reionization & Near-Field Cosmology

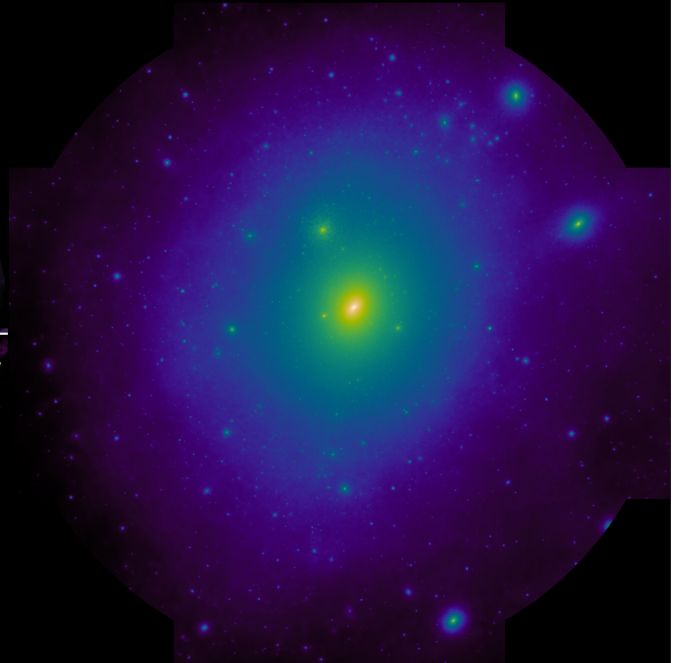


FIG. 1.— Summary of the four key science drivers of the MEGATRON simulations. **Top left:** A $z \sim 15$ density map with $H\beta$ (white) and $[O\ III]\ \lambda 5007$ (yellow) shows radiation-driven $H\beta$ emission extending into the IGM, while metals remain near galaxies. **Top right:** A rotationally-supported $z = 10$ galaxy reveals early H_2 and CO formation and complex emission line morphology. **Bottom left:** The CGM of a massive $z \sim 4$ star-forming galaxy highlights how different ions trace distinct gas phases. **Bottom right:** The $z = 0$ dark matter distribution connects high- and low-redshift structure formation.

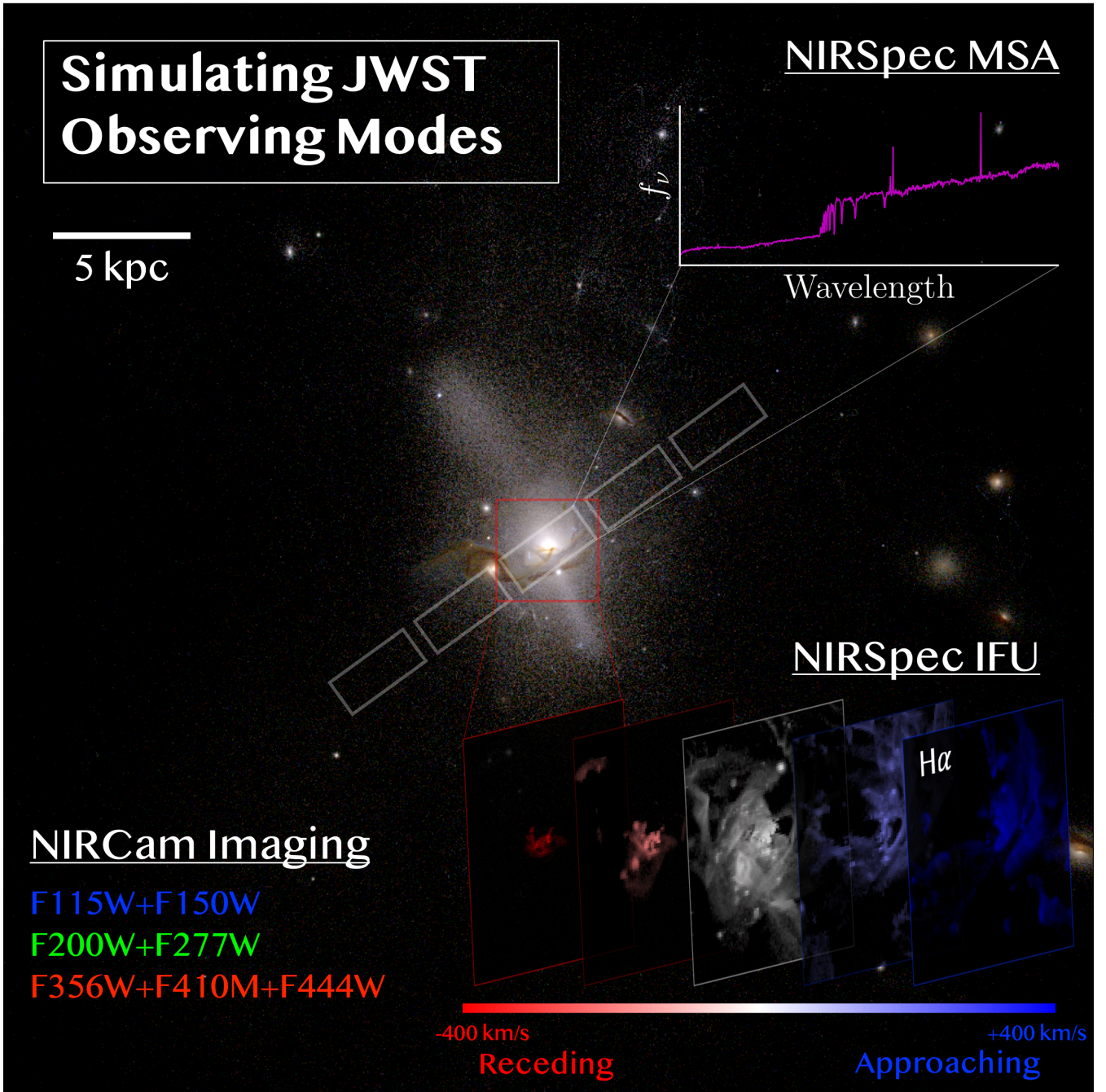


FIG. 2.— Demonstration of how MEGATRON galaxies can be mock observed in multiple JWST observing modes. The example galaxy is a dusty spiral at $z = 4$ from the Cosmic Noon suite. The background RGB image combines JWST NIRCam filters F115W and F150W, in the blue channel, F200W and F277W, in the green channel, and F356W, F410M, and F444W in the red channel. Prominent dust lanes are visible as are large plumes of stars that are a remnant of a merger. To mock the NIRSpec micro shutter array (MSA), we overlay slitlets and show the spectrum of the central region of the galaxy. The object has a very red UV slope from dust attenuation and a clear Balmer break due to an aging stellar population, and some weak emission lines that are remnants of previous and ongoing star formation. Finally, we mock the central region of the NIRSpec IFU focusing on the velocity-resolved H α line. We split the velocity channels in to bins of $v < -400$ km/s, $-400 < v < -200$ km/s, $-200 < v < 200$ km/s, $200 < v < 400$ km/s, and $v > 400$ km/s. Note how the morphology of the H α line strongly depends on velocity. All images, spectra, and IFU data were created with the Monte Carlo radiation transfer code RASCAS (Michel-Dansac et al. 2020).

inputs are loosely motivated by various physical mechanisms that may help explain the excess numbers of bright galaxies at high redshift (e.g., Chemerynska et al. 2023; Leung et al. 2023; Harikane et al. 2024; Finkelstein et al. 2023), while achieving high-enough resolution to predict the spectral properties of the ISM that can be directly compared to existing JWST and ALMA data. Three simulations run at slightly lower resolution are designed to capture the non-equilibrium physics in the CGM towards cosmic noon, with the aim of studying the observable properties of galaxies that can be used to probe the cosmic baryon cycle. The CGM simulations will be presented in detail in a companion paper (Cadiou et al. 2025).

Details of the seven simulations can be found in Table 1 and a summary of the methods employed for galaxy formation physics, halo finding, and dust radiation transport are described in Appendix A. In this work, we focus primarily on the high-redshift simulation suite which is stopped at $z \sim 8.5$, while the cosmic noon simulations (which end closer to $z \sim 3$) is presented in Cadiou et al. (2025). Furthermore, the Milky Way mass initial conditions of these simulations are designed to allow us to make connections with Milky Way satellite population in the low-redshift Universe and the properties of Milky Way progenitors. A dark matter-only simulation using the same initial conditions has been run to $z = 0$ and the connection between high and low redshift is presented in a second companion paper (Rey et al. 2025).

2.1. Initial Conditions

We construct cosmological, zoomed initial conditions (ICs) for a Milky-Way-mass halo (with virial mass of $\approx 10^{12} M_\odot$ at $z = 0$). This halo is based on a reference object first presented in Rey & Starkenburg (2022) and evolved to $z = 0$ for the VINTERGATAN-GM project (Agertz et al. 2021; Rey et al. 2023).

We generate ICs using the genetIC software (Stopyra et al. 2021) and a flat Planck Collaboration et al. (2016) cosmology with $\Omega_M = 0.3139$, $h = 0.6727$, $\sigma_8 = 0.8440$, $n_s = 0.9645$ and $\Omega_b = 0.04916$. From a dark matter-only cosmological volume with a box size of $50 \text{ Mpc } h^{-1} = 73 \text{ Mpc}$ and mass resolution $m_{\text{DM}} = 1.2 \times 10^8 M_\odot$, we identify an isolated halo with Milky-Way virial mass ($M_{200} \approx 10^{12} M_\odot$ and no more massive neighbors within $5r_{200c}$, where r_{200c} is the radius enclosing 200 times the critical density of the Universe; see Rey & Starkenburg 2022 for further details).

We trace the region enclosing $3r_{200c}$ at $z = 0$ around this object to the low-resolution ICs and refine the mass resolution within this Lagrange region to $m_{\text{DM}} = 1.67 \times 10^4 M_\odot h^{-1}$ (effective resolution 8912^3). Importantly, since this region is selected from the $z = 0$ halo, it represents a large cosmic volume at high redshift (initially $18 \text{ comoving Mpc}^3 h^{-3}$) and ensures several thousands of high-redshift galaxies are adequately resolved without contamination.

The mass growth history of our reference object is shown as the blue dashed line in Figure 3. To further maximize the size of the high-redshift galaxy population, we use a quadratic genetic modification technique, as outlined in Rey & Pontzen (2018); Rey, Pontzen & Sainlonge (2019), to modify the collapse of progenitors within

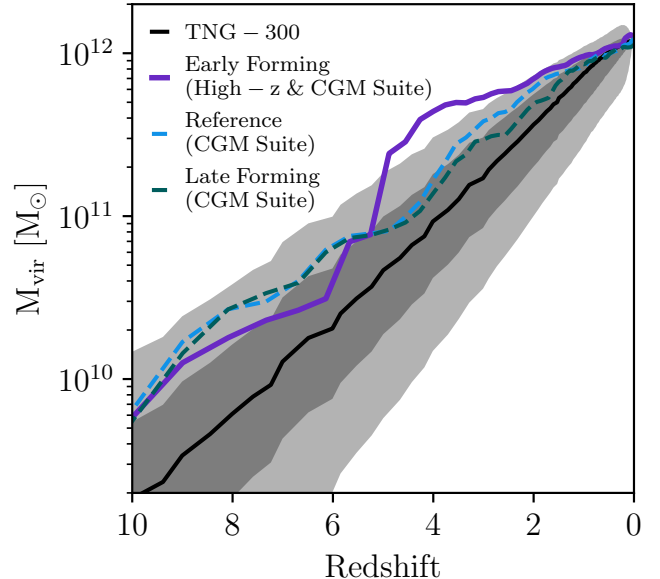


FIG. 3.— Mass growth histories of the main progenitor in the three sets of initial conditions compared to the typical mass growth histories of 28,475 similar mass haloes from IllustrisTNG-300. The dark and light shaded regions represent the 1σ and 2σ results from IllustrisTNG-300 with the black line representing the median relation. The early forming halo (purple) is used for the high-redshift suite as it results in more numerous massive progenitors at high redshift rather than a single dominant object. The reference and late forming initial conditions (dashed blue and green lines, respectively) are only used for the CGM suite.

the Lagrangian patch. More specifically, we increase the variance by 10% on scales of $0.30 \text{ Mpc } h^{-1}$. This ensures that all progenitor haloes with masses $\approx 2 \times 10^{10} M_\odot$ collapse earlier, generating a larger sample of more massive high-redshift galaxies, which is useful for comparing with JWST data. The result of this modification is to initially delay the mass accretion onto the main progenitor (purple line in Figure 3) since mass is spread across multiple, more massive progenitors early on. However, the modification promotes a rapid assembly at $z \sim 6 - 4$ when the multiple, more massive progenitors merge. To ensure a fixed total mass at $z = 0$, we reduce the mean overdensity of the Lagrangian region by 12% (see also Roth, Pontzen & Peiris 2016; Pontzen et al. 2017 for further details).

Finally to study the impact of the timescale of the collapse of the Lagrange region on the CGM towards cosmic noon, we create a third IC (green dashed line in Figure 3) that assembles slightly faster at early times and is delayed at later times by reducing the variance by 5% on scales of $0.30 \text{ Mpc } h^{-1}$.

To assess the likelihood of the formation scenarios we engineered with the genetic modification approach, we first compute the likelihood difference between the reference and modified initial conditions. We find $\Delta\chi^2 = -2630$ and $\Delta\chi^2 = 9530$ for the ‘Early’ and ‘Late’ ICs, respectively. Both $\Delta\chi^2$ are low compared to the number of degrees of freedom available in our Lagrangian region which contains $\approx 10^8$ modes, ensuring that our ICs are compatible with the Λ CDM power spectrum. We also compare in Figure 3 our mass assembly histories to a

TABLE 1

DETAILS OF THE SIMULATIONS IN THE MEGATRON SUITE. WE LIST THE SIMULATION NAMES, THE SUITE THEY BELONG TO, THE MASS OF THE DARK MATTER PARTICLES, MINIMUM MASS OF THE STELLAR PARTICLES, MINIMUM PHYSICAL CELL SIZE AT $z = 8.5$ OR THE TARGET CONSTANT PHYSICAL RESOLUTION, INITIAL METALLICITY, THE CURRENT REDSHIFT OF THE SIMULATION, THE VIRIAL MASS OF THE PRIMARY HALO AT THE CURRENT REDSHIFT, THE JEANS LENGTH BELOW WHICH STAR FORMATION IS TRIGGERED, THE EFFICIENCY OF STAR FORMATION PER FREE-FALL TIME, THE ENERGY PER SN, AND WHETHER THE SIMULATION INCLUDED HYPERNOVAE (HN). A STAR FORMATION EFFICIENCY DENOTED AS $f(\sigma_V)$ INDICATES THAT THE MODEL USES A VARIABLE EFFICIENCY RELATED TO THE LOCAL TURBULENT PROPERTIES OF THE GAS.

Name	Suite	m_{DM} ($M_\odot h^{-1}$)	m_* ($M_\odot$)	Δx_{min} (pc h^{-1})	Z_{initial} $Z_\odot$	z_{current}	$M_{\text{vir,max}}$ $\log_{10}(M_\odot)$	$\lambda_{\text{J,SF}}$	ϵ_{ff}	E_{SN} (10^{51} erg)	HN
Efficient SF	High-Redshift	1.67×10^4	500	~ 5	0	8.5	10.30	Δx	$f(\sigma_V)$	1.0	$\times$
Bursty SF	High-Redshift	1.67×10^4	500	~ 5	0	8.5	10.27	Δx	$f(\sigma_V)$	5.0	$\times$
Variable IMF	High-Redshift	1.67×10^4	500	~ 5	0	8.5	10.30	Δx	$f(\sigma_V)$	1.0	$\checkmark$
HN, High ϵ_{ff}	High-Redshift	1.67×10^4	2,000	~ 5	0	8.5	10.30	$4\Delta x$	100%	1.0	$\checkmark$
Early Collapse	Cosmic Noon	1.67×10^4	4,600	~ 20	10^{-4}	2.94	11.67	Δx	$f(\sigma_V)$	1.0	$\times$
Fiducial Collapse	Cosmic Noon	1.67×10^4	4,600	~ 20	10^{-4}	3.60	11.26	Δx	$f(\sigma_V)$	1.0	$\times$
Late Collapse	Cosmic Noon	1.67×10^4	4,600	~ 20	10^{-4}	3.32	11.29	Δx	$f(\sigma_V)$	1.0	$\times$

large sample of $\approx 30,000$ mass assembly histories for haloes with a similar $z = 0$ mass in the Illustris TNG-300 (Pillepich et al. 2018; Nelson et al. 2018, see Rey & Starkenburg 2022 for more details of the comparison). All of our ICs assemble more rapidly $z \geq 4$ compared to a typical Milky-Way-mass halo, with the ‘Early Forming’ IC used for the high-redshift suite having an excursion above the 2σ contour at intermediate redshift ($z \sim 5-3$).

We emphasize that our chosen ICs are not typical of haloes of similar mass at $z = 0$, but rather engineered for specific purposes — having a Milky Way mass galaxy at $z = 0$ while still producing a significant number of high-redshift galaxies. While this is ideal for our science goals focused on studying the internal properties of galaxies (e.g., the structure of the ISM and CGM, how different modes of star formation are realized in observable properties, etc.), comparing with population statistics (e.g., luminosity functions, stellar mass functions, etc.) will require more care given the biased growth history.

All ICs are evolved to $z = 149$ using first-order linear theory (Zel’dovich 1970). For the high-redshift suite simulations, we assume a primordial composition of 76 % hydrogen and 24 % helium, and no metals. For the cosmic noon suite, we adopt a metallicity of $12 + \log_{10}(\text{O}/\text{H}) = 4.69$ ($10^{-4} Z_\odot$) for the oxygen abundance and follow Rammason et al. (2022) to compute the abundances of other elements⁴.

To reduce advection errors and increase the accuracy of high-redshift hydrodynamics, we ‘genetically’ modify the initial conditions to ensure that the Lagrangian region is at rest compared to the numerical grid following the procedure described in Pontzen et al. (2021).

2.2. Summary of Simulations

The key differences between the seven simulations are summarized in this section. The high-redshift suite uses a constant comoving resolution which results in extremely high-resolution (up to ~ 1 pc physical resolution at the time when the first Pop. III stars form) and is thus better suited for studying the details of star formation and the ISM at high redshift. The cosmic noon suite employs a constant physical resolution of ≈ 20 pc h^{-1} . This is computationally less expensive, allowing us to run to lower redshift. However, the sacrifice in spatial reso-

⁴ Note that we assume a solar metallicity of $12 + \log_{10}(\text{O}/\text{H}) = 8.69$ (Asplund et al. 2009).

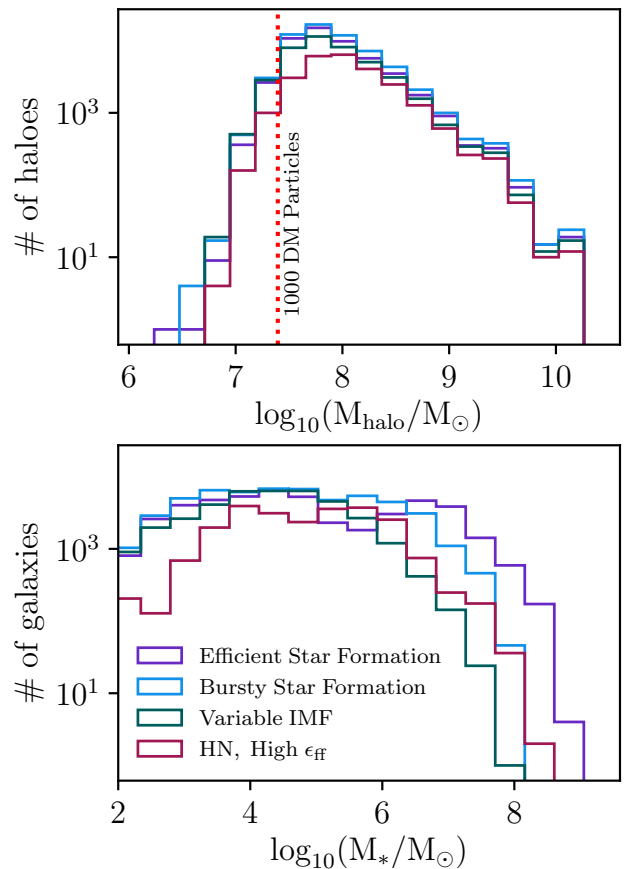


FIG. 4.— Histograms of halo mass (top) and stellar mass (bottom) across all snapshots for each simulation in the high-redshift suite. In the top panel we indicate the halo mass that corresponds to 1,000 DM particles, which represents the minimum required to be included in our spectroscopic sample.

lution makes these simulations more ideal for studying more massive galaxies and the CGM.

Within the high-redshift suite, we consider variations to feedback, IMF, and star formation. In the Bursty star formation simulation, the energy injected per SNII is increased to 5×10^{51} erg, which results in large fluctuations in galaxy star formation rates. The variable IMF simulation allows for hypernova and a density and metallicity-dependent IMF. Finally, in the HN, High ϵ_{ff}

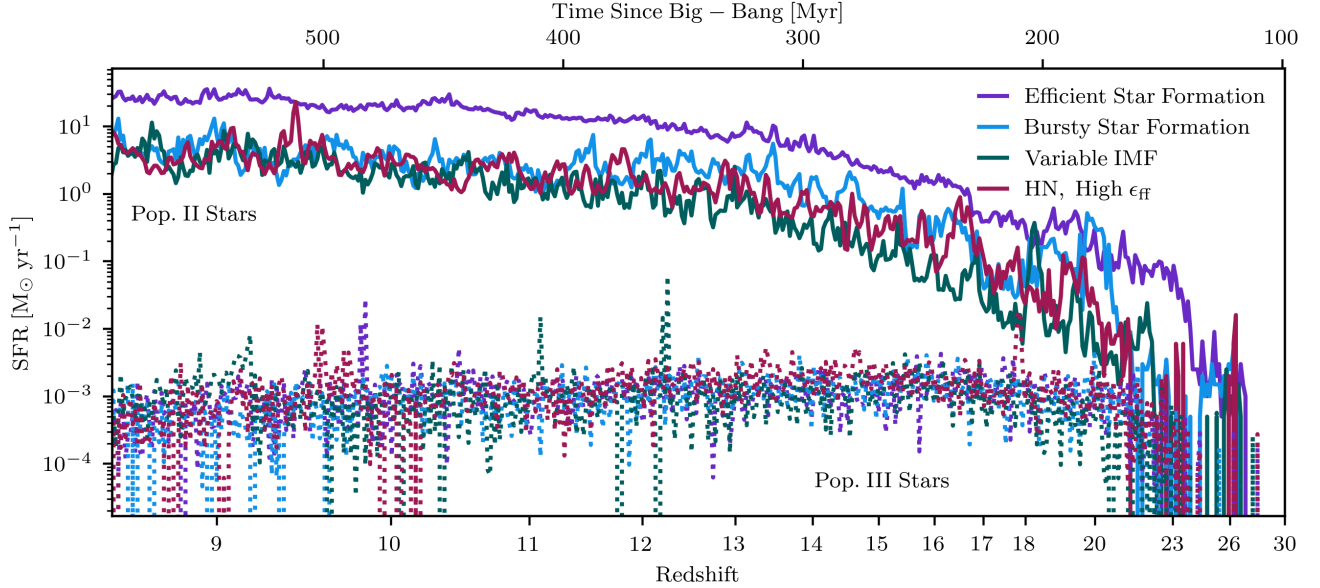


FIG. 5.— Star formation histories of Pop. II stars (solid) and Pop. III stars (dotted) across the entire Lagrange volume for each of the high-redshift suite simulations.

simulation, we require the Jeans length to be better resolved for star formation to occur, form stars at 100% efficiency per free-fall time, and allow for hypernova.

The galaxy formation physics in the cosmic noon suite is identical to the efficient star formation model as this is closest to what was used in VINTERGATAN. The only difference between the three cosmic noon simulations is the ICs. These ICs are systematically designed to study the impact of an earlier and later formation history on the properties of the lower-redshift CGM.

For all high-redshift simulations, we output full simulation snapshots at a cadence of 5 Myr after $z = 30$ ⁵. This cadence is increased to 20 Myr for the CGM suite.

3. STAR & GALAXY FORMATION IN A MILKY WAY MASS LAGRANGIAN REGION

We begin by describing the basic physical properties of the dark matter, gas, and stars in the Lagrange region as a function of time for the simulations that are part of the high-redshift suite. The top panel of Figure 4 shows the distribution of dark matter halo masses in our simulations, combining results from all snapshots up to $z = 8.5$. By genetically modifying the reference ICs, we resolve thousands of haloes in the epoch of reionization, up to masses $> 10^{10} M_{\odot}$, as well as those well below the atomic cooling threshold, leading to a statistical sample size that is ideal for comparing to galaxies at the faint-end of the UV luminosity function at high redshift. Across all snapshots, considering only those haloes with at least one star particle and resolved by $> 1,000$ dark matter particles and $> 1,000$ gas cells, which meet our criteria to be included in our spectroscopic sample, the simulations produce a data set of $> 175,000$ spectra. This represents a nearly $100\times$ increase compared to the

number of individual galaxies with spectra computed for SPHINX²⁰ (Katz et al. 2023b) and RENAISSANCE (Barrow et al. 2017) simulations, which are representative of those that have a resolved ISM and on-the-fly radiation transport.

Star formation first occurs in the simulations at $z \sim 27$. In our model, a single Pop. III explosion, regardless of mass, can provide enough heavy elements to reach the critical metallicity for Pop. II star formation (see also e.g., Wise et al. 2012; Brauer et al. 2025). By $z = 26$, Pop. II star formation is already dominant across the Lagrange region. This is demonstrated in Figure 5 where we show the star formation history of the zoom region for all high-redshift simulations. By $z = 20$ the Pop. II SFR is typically an order of magnitude or more greater than that of Pop. III stars. The effect is strongest when weaker feedback allows more efficient Pop. II star formation (purple). Despite the dominance of Pop. II star formation in all models at $z \lesssim 26$, a perhaps surprising result is that Pop. III stars are still forming in the simulation at $z = 8.5$. Late Pop. III star formation has been observed in previous simulations (e.g., Pallottini et al. 2014; Xu et al. 2016; Jaacks et al. 2018; Sarmiento, Scannapieco & Cohen 2018), although this typically occurs in low-density regions, and not necessarily in the Lagrange region around a Milky Way mass object.

The bottom panel of Figure 4 shows a histogram of the stellar masses of the galaxies in our spectroscopic sample (after accounting for mass loss from stellar evolutionary processes). Depending on the simulations, we find stellar masses up to $10^8 - 10^9 M_{\odot}$. The efficient star formation model produces the highest stellar masses, again due to the weaker feedback, while in the variable IMF model, the galaxy stellar masses are much better regulated. The stellar masses are better controlled in the variable IMF model partially because the energetic feedback budget is increased due to the inclusion of HN and the increase in

⁵ Note that we store many more snapshots at irregular time intervals, as the simulation is checkpointed just before reaching the wall-clock time limit.

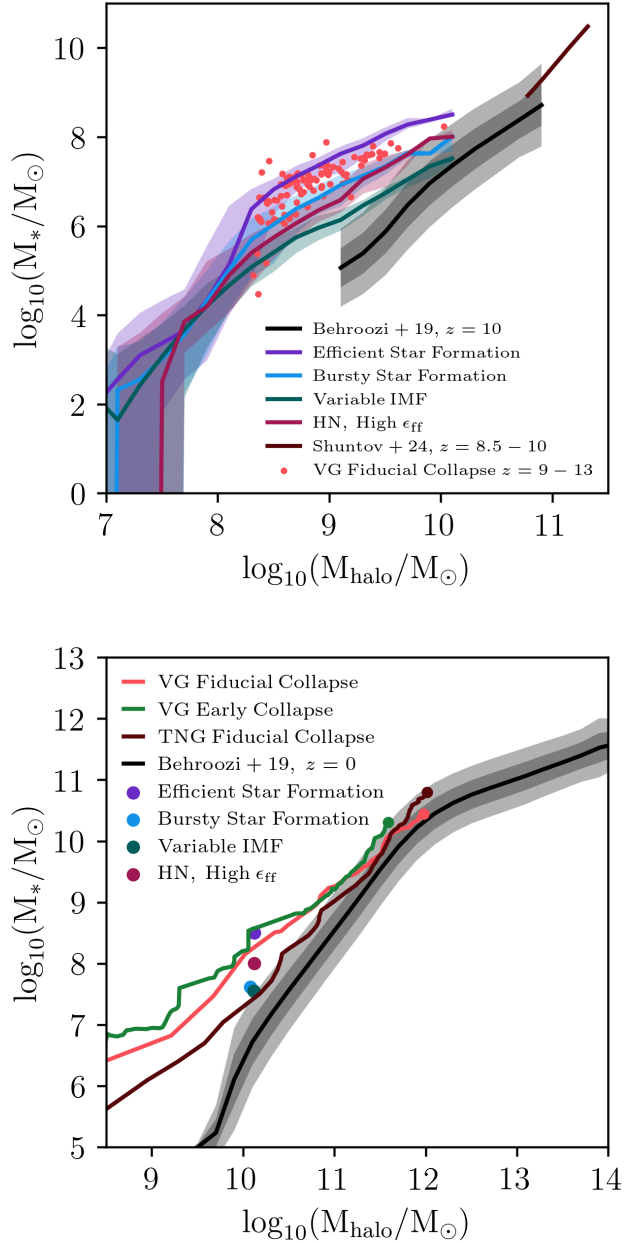


FIG. 6.— (Top) Stellar mass - halo mass relation for each simulation, including all galaxies at every redshift. We show the median and 1σ scatter about the relation. For comparison, the empirical model of Behroozi et al. (2019) at $z = 10$ (black) and its 1 and 2σ scatter as well as inferences from high-redshift JWST observations from Shuntov et al. (2025) (brown line). We also show the results from the VINTERGATAN simulation (salmon points) of the fiducial collapse initial conditions, which serve as our benchmark. (Bottom) Stellar mass - halo mass relation for VINTERGATAN simulations of the fiducial (salmon line) and early collapse (green line) compared with the $z = 0$ empirical constraints from Behroozi et al. (2019) at $z = 0$ (black). Note that there are more simulation outputs for the early collapse model, which explains why the line is less smooth and the simulation has not been run to $z = 0$. The brown line shows the fiducial initial conditions run with the Illustris-TNG model Joshi et al. (2025). Here we see that the benchmark halo is in good agreement with empirical constraints at low redshift, despite being a more than 2σ outlier at $z \sim 9 - 13$. Coloured points represent the values for the most massive progenitor halo at $z = 8.5$ in each of the high-redshift suite simulations.

number of SN due to a shallower upper-mass IMF slope. In addition, the increased number of massive stars in the variable IMF model leads to the star particles losing a greater fraction of their initial mass. As can be seen in Figure 5, by $z = 8.5$, the SFR in the Lagrange region is similar in the bursty star formation and variable IMF models and significantly lower than in the efficient star formation simulation. However, the extra mass loss from the top-heavy IMF reduces the stellar mass of the variable IMF simulation with respect to the bursty star formation model.

The regulation of star formation is explored in the top panel of Figure 6 where we show the stellar mass-halo mass relation for all high-redshift simulations. The benchmark for our simulation are the results from VINTERGATAN (shown as salmon points) for the set of initial conditions with the fiducial collapse. The VINTERGATAN simulations have been run to $z = 0$ and good agreement was found with the local stellar mass-halo mass relation (see bottom panel of Figure 6 and also Agertz et al. 2021). The efficient star formation model uses a feedback scheme very similar to VINTERGATAN (by design) and we find a small offset such that this MEGATRON model forms $\sim 50\%$ more stars compared to our benchmark. This is partially due to the initial condition modification for early collapse but also likely caused by the higher DM resolution, the adoption of constant physical versus constant comoving resolution, and the non-equilibrium cooling used in MEGATRON. Note in the bottom panel of Figure 6 that the stellar mass-halo mass relation for the main progenitor in the early collapse initial conditions is consistently higher than for the fiducial collapse. Nevertheless, the agreement between our simulations and VINTERGATAN is very acceptable given the minor changes. The other three models form fewer stars than the benchmark VINTERGATAN simulation due to the enhanced feedback, with the variable IMF simulation having the lowest stellar masses. At $z = 8.5$, our four high-redshift simulations bracket the stellar masses expected for the same set of early collapsing initial conditions from the PARADIGM project (Joshi et al. 2025). In that work, the early collapse initial conditions were simulated with both the VINTERGATAN and Illustris-TNG models. Interestingly, they found that the Illustris-TNG model forms fewer stars at high redshift compared to the VINTERGATAN model but then falls higher on the stellar mass-halo mass relation at $z = 0$. Our results thus span reasonable predictions of the high redshift Universe based on models calibrated at $z = 0$.

For further comparison, we show the inferred stellar mass-halo mass relation from recent JWST survey data at high redshift (Shuntov et al. 2025), which unfortunately does not overlap in mass range with our models as well as empirical constraints from Behroozi et al. (2019). We caution comparisons with Behroozi et al. (2019) because this model is known to significantly underpredict galaxy number counts at high redshift, even at $z < 9$ (e.g. Finkelstein et al. 2023). Nevertheless, this comparison contextualizes our results. From this figure, we emphasize two key points: 1) as discussed in Katz et al. (2024c), agreement with the stellar mass-halo mass relation at $z = 0$ does not imply significant regulation at high redshift. Most of the progenitor haloes in our model collapse into a single object at $z = 0$ and thus

the mild offset we see for the efficient star formation run compared to the benchmark is not concerning. In fact, if we were able to run to $z = 0$, the three models with stronger feedback may over-regulate the galaxy because the benchmark halo falls well within the 1σ scatter of the Behroozi et al. (2019) empirical model at $z = 0$. 2) Our initial conditions represent a highly biased environment due to the engineered formation history. Compared to the other VINTERGATAN haloes, this particular object has a higher conversion rate of gas into stars at high redshift due to the early collapse. Again, our results do not preclude agreement with the $z = 0$ stellar mass-halo mass relation (as shown in the bottom panel of Figure 6).

Because our simulated Lagrange region represents that of a Milky Way mass halo, one particularly interesting question is when the progenitors first become observable. In Figure 7, we show the UV magnitudes of bright resolved galaxies as a function of redshift for all high-redshift suite simulations. Note that here and throughout this work, the UV magnitude represents the intrinsic value at $1500 \text{ \AA}$, considering both the stellar and nebular continuum⁶. For context, we also show spectroscopically confirmed high-redshift galaxies from the JADES survey (Carniani et al. 2024; Bunker et al. 2024; D’Eugenio et al. 2025) as well as the photometrically selected high-redshift galaxy candidates from the GLIMPSE survey (Kokorev et al. 2024; Chemerynska et al. 2025), and by Castellano et al. (2025). There is significant overlap in M_{UV} -redshift space between the high-redshift photometric candidates from GLIMPSE and the simulated MEGATRON galaxies. All simulations except for the HN, High ϵ_{ff} run predict that certain Milky Way progenitors can reach UV magnitudes brighter than -17 at $z > 17$, which is now within reach of the deepest JWST imaging surveys around lensing clusters. Thus, we may be able to detect galaxies that evolve into those like our own at a redshift representing $\sim 1.5\%$ of the current age of the Universe. While this is extremely promising for the opportunity to constrain the origins of our own galaxy, it is clear from Figure 7 that JWST will only detect the absolute brightest progenitors that represent the tail end of the distribution for what is typical in such an environment at these redshifts.

4. THE SPECTRAL DIVERSITY OF HIGH-REDSHIFT GALAXIES

A key aspect of MEGATRON is the ability to predict the intrinsic spectra of galaxies. Across all four high-redshift simulations, we produce $> 175,000$ angle-averaged spectra, accounting for the stellar continuum, nebular line emission, and nebular continuum emission. In this work, we ignore the role of dust absorption, re-emission, and scattering, all of which will be considered in subsequent work (see Appendix A). This is likely a reasonable assumption since we focus primarily on low-mass objects and most of the observed high-redshift galaxies exhibit β slopes that are suggestive of minimal dust attenuation⁷ (e.g., Cullen et al. 2024; Top-

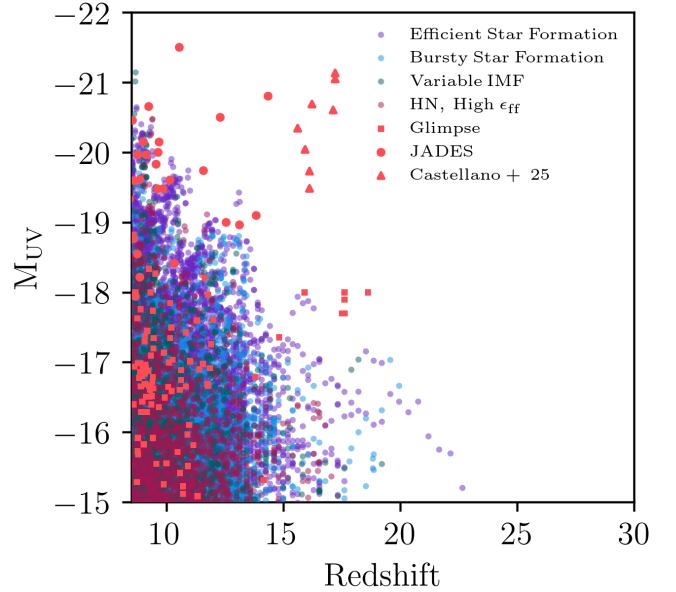


FIG. 7.— Intrinsic UV magnitude as a function of redshift for galaxies in the high-redshift suite simulations. For context, we also show spectroscopically confirmed high-redshift galaxies from the JADES survey (Carniani et al. 2024; Bunker et al. 2024; D’Eugenio et al. 2025) as well as the photometrically selected high-redshift candidate galaxies from the GLIMPSE survey (Kokorev et al. 2024; Chemerynska et al. 2025) and by Castellano et al. (2025). Note that we have excluded uncertainties on the photometric data points for clarity.

ping et al. 2024; Saxena et al. 2024).

To qualitatively demonstrate the types of spectra that naturally emerge from our simulations, we perform dimensionality reduction and compute a 2D embedding for all $> 175,000$ spectra in the suite using UMAP⁸ (McInnes et al. 2018; McInnes, Healy & Melville 2018). This exercise aims to reveal the variety of spectral types in our simulations by reducing their high-dimensional representations to a 2D plane, preserving both local relationships and global structure. Spectra are all normalized to their value at rest-frame $3,000 \text{ \AA}$ in f_ν and embeddings are computed in the rest-frame. In Figure 8, we show this 2D embedding and plot the median stacked spectra of visually selected regions. The types of galaxies we identify are:

1. Pop. III galaxies and cooling haloes that have strong emission H and He emission lines, no (or very weak) metal lines, and a dominant nebular continuum.
2. Mini-quenched galaxies with very weak or no emission lines and a strong Balmer break.
3. Post-starburst galaxies with weak emission lines and a Balmer break.
4. Starforming galaxies with a blue UV continuum, and strong H, He, and metal emission lines.

⁶ For such low mass galaxies, dust attenuation is not expected to be significant (e.g. Ma et al. 2018; Rosdahl et al. 2018), but it will be considered in future work.

⁷ Although blue UV slopes could also be due to a gray attenuation curve (e.g. McKinney et al. 2025).

⁸ We set the number of neighbours to 100, and a minimum distance of 0 for the UMAP decomposition.

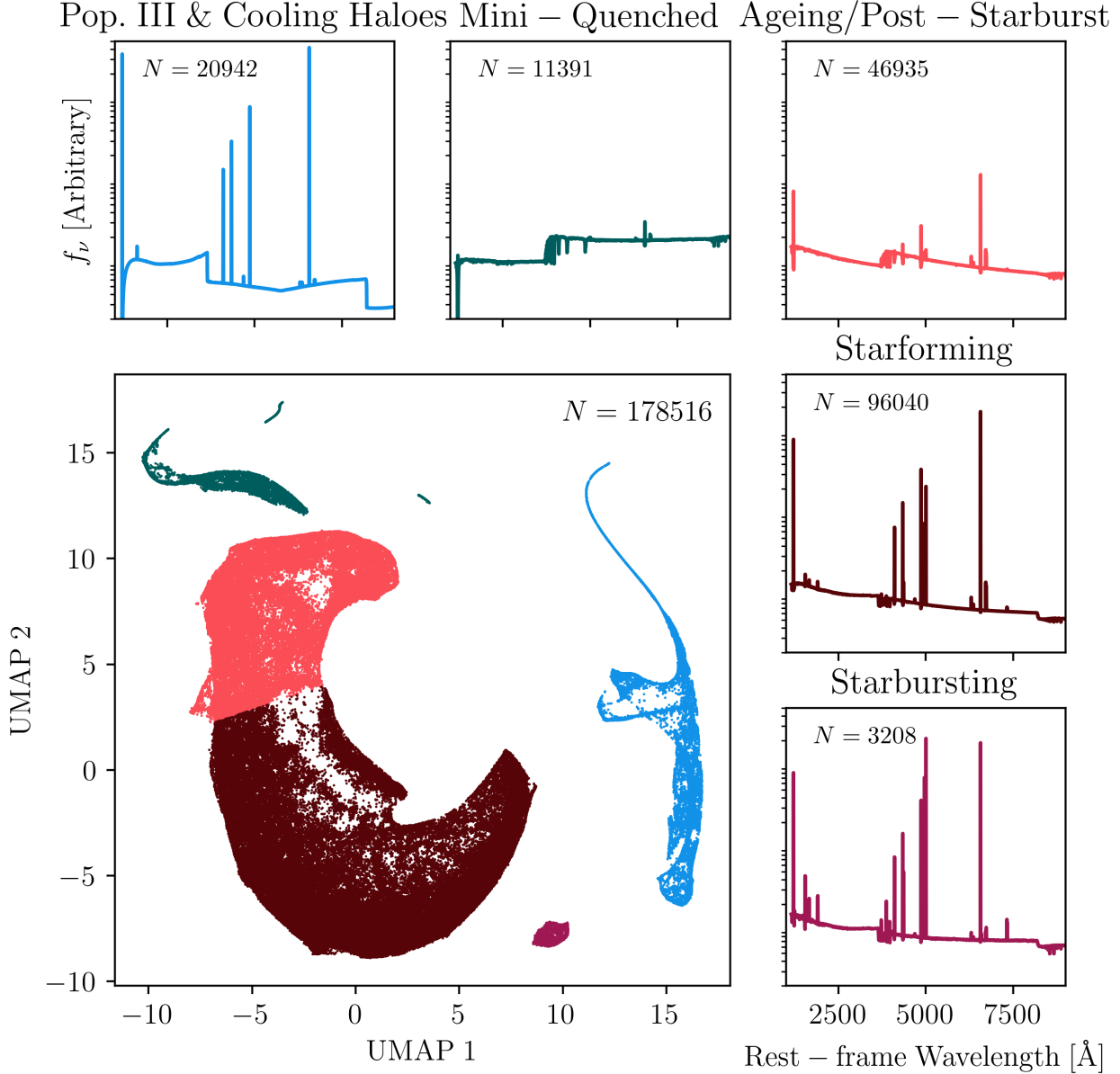


FIG. 8.— UMAP embeddings for all 178,516 spectra from the high-redshift suite. For each visually selected sample of points, we show the median stacked spectrum and identify the galaxy type. The number of galaxies in each sample is listed in each panel.

5. Starbursting galaxies with extremely strong H, He, and metal emission lines, a blue UV continuum, and a Balmer jump.

The majority of galaxies in our simulations are normal starforming galaxies. These objects form a smooth continuum with the aging/post-starburst galaxies that represent the next most populous group. Various clusters of objects separate themselves from this connected continuum due to their peculiar spectral properties. In this section, we explore this diversity of spectral types in detail and place them in the context of recent JWST observations.

4.1. Population III Galaxies

The first, clearly-isolated cluster in Figure 8 maps onto the spectra of galaxies dominated by Pop. III stars (blue). This includes haloes with both live Pop. III star formation and haloes that are cooling after the Pop. III stars have died. The lack of metal lines, strong He II emission, and dominant nebular continuum are clear signatures of Pop. III physics.

The UV magnitudes of Pop. III galaxies in our simulations tend to be very faint. In the top panel of Figure 9, we show a histogram of 1500 $\text{\AA}$ UV magnitudes of starforming Pop. III galaxies. The vast majority of Pop. III stars in the simulations form in haloes close to the atomic cooling limit, either individually or in small groups, lead-

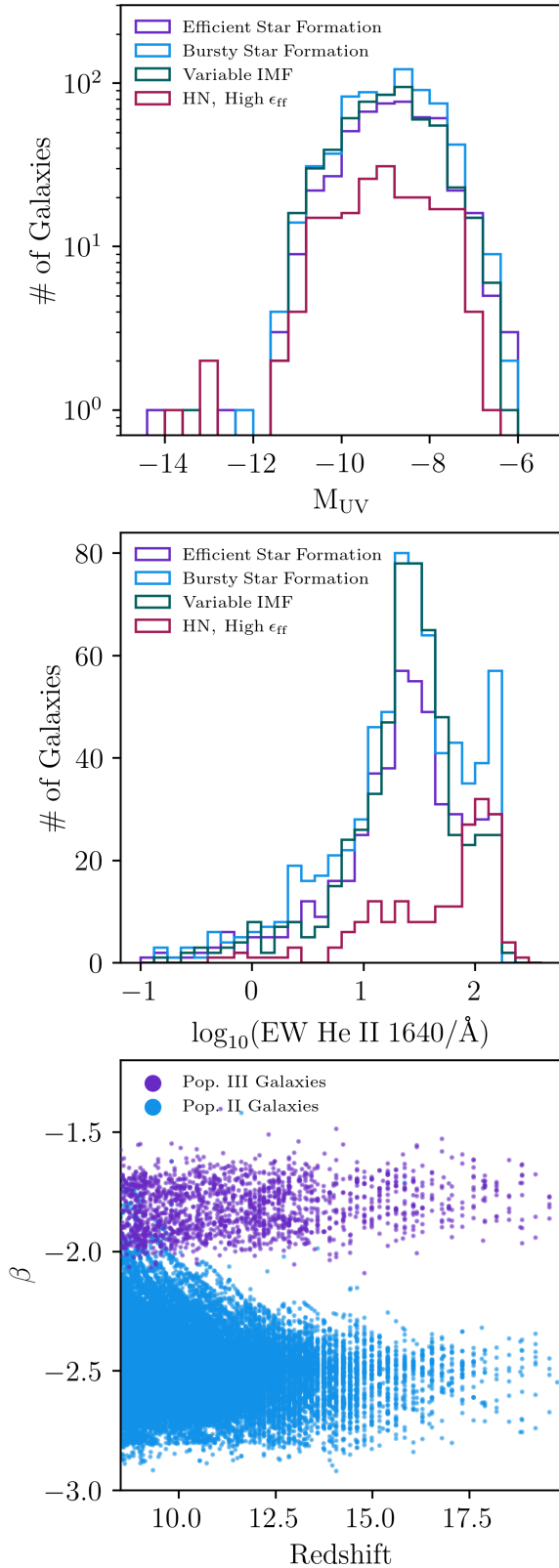


FIG. 9.— (Top) Histogram of the 1500 Å magnitude of Pop. III galaxies. (Middle) Histogram of the He II EW of Pop. III galaxies. (Bottom) Intrinsic UV slopes of Pop. III (purple) and Pop. II (blue) galaxies as a function of redshift. We only consider galaxies where Pop. III stars have formed in the previous 1 Myr.

ing to UV magnitudes fainter than -11 . Objects of this magnitude have yet to be observed at high redshift, with the faintest strongly-lensed object in the reionization era having an M_{UV} of -12.3 (Vanzella et al. 2024; Nakajima et al. 2025).

Among all of the simulations, there are seven Pop. III galaxies that break the $M_{UV} < -12$ barrier, reaching magnitudes as bright as -14.4 , well within the reach of JWST surveys leveraging gravitational lensing (e.g. Fujimoto et al. 2025). This is highly encouraging for the prospects of detection, even if such systems are extremely rare (7/20,942). Storck et al. *in prep.* will explore in detail the physical conditions required to produce bright Pop. III galaxies and their sensitivity to galaxy formation assumptions.

However, differentiating Pop. III galaxies from their low-metallicity Pop. II counterparts remains a key concern for robustly detecting them. This is especially true at high redshifts (e.g., at $z \gtrsim 10$) where strong rest-frame optical metal lines drop out of the NIRSpect PRISM and one must rely on lower equivalent width UV lines. He II 1640 Å has long-been considered one of the key signatures of massive Pop. III stars (e.g., Tumlinson & Shull 2000; Oh, Haiman & Rees 2001; Schaerer 2002), but as we show in the middle panel of Figure 9, not all Pop. III galaxies in MEGATRON have high He II EWs. A lack of He II 1640 Å does not imply that a galaxy is not a Pop. III system. This is a direct result of the fact that not all Pop. III stars are expected to be massive and have extreme surface temperatures $\gtrsim 10^5$ K (e.g., Clark et al. 2011; Hirano et al. 2014; Stacy, Bromm & Lee 2016). It should also be noted that strong, narrow He II 1640 Å can also be generated by sources other than Pop. III stars, for example AGN or Wolf-Rayet stars with weak winds (e.g., Gräfener & Vink 2015), and thus this particular diagnostic is not necessarily definitive.

One of the key differentiating factors between Pop. III and Pop. II galaxies that we find in the simulations is that Pop. III galaxies are generically redder. This is counterintuitive because the intrinsic UV slopes of massive Pop. III stars are extremely blue, with values much steeper than -3 (e.g., Schaerer 2002; Tumlinson, Shull & Venkatesan 2003; Larkin, Gerasimov & Burgasser 2023). However, Pop. III galaxies represent a rare class of systems where the nebular continuum can outshine the stars in the UV which leads to redder slopes. This has been demonstrated previously using photoionization models (e.g., Raiter, Schaerer & Fosbury 2010; Trussler et al. 2023; Katz et al. 2024a), where the effect highly depends on the electron density of the nebula due to the relatively low critical density of H I two-photon emission. As we show in the bottom panel of Figure 9, the UV slopes of Pop. III galaxies are typically $\gtrsim -2$, which is much redder than the standard Pop. II starforming galaxies in MEGATRON at the same redshift. In fact, the intrinsic Pop. II galaxy slopes only become this red at lower redshifts when the stars have aged enough. For this reason, photometric selection cuts for high-redshift galaxies should not exclude systems with UV slopes $\gtrsim -2$.

For the same reason that Pop. III galaxies appear redder than Pop. II galaxies — i.e. due to their strong nebular continuum emission — they also exhibit a downturn in the UV, blueward of Ly α , which is a signa-

ture of intense two-photon emission. Such a feature has rarely been discussed in the literature (although see [Dijkstra 2009](#); [Raiter, Schaerer & Fosbury 2010](#); [Katz et al. 2024a](#)), but has gained prominence more recently due to its possible detection at high redshift ([Cameron et al. 2024](#); [Katz et al. 2024a](#); [Witstok et al. 2024](#), albeit with significant metal line emission).

Many of the spectral features discussed in this section can be seen in Figure 10 where we show the spectra of seven example bright Pop. III galaxies. These systems all show strong He II emission in the UV and optical, a Balmer (and Paschen) jump, and a downturn in the UV from two-photon emission. The different strengths of the UV downturn depend on the electron density near the star particles and the fact that the spectra appear flat or slightly rising in f_ν indicates that the spectral slopes are redder than ~ -2 .

If the energy injection from the Pop. III radiation and SNe is high enough, star formation can be quenched and cooling radiation from the hot gas may also be detectable. This has been considered previously in the more general case of cooling clouds (e.g., [Dijkstra 2009](#)) and also in the case of Pop. III remnant galaxies ([Katz et al. 2023a](#)). In Figure 11, we show a stacked spectrum of $> 10,000$ galaxies (across all high-redshift simulations) that are cooling after having a Pop. III star formation event (pure line emitters). Their spectra are characterized by having no stellar emission and are thus purely nebular line and continuum emitters. This means that the equivalent widths of the lines are extremely strong (thousands of Å for H α and H β depending on temperature, see Appendix A of [Katz et al. 2024a](#)) and the nebular two-photon and Balmer/Paschen jump features are highly visible. The strength of the Balmer/Paschen jump and the line EWs provides a strong constraint on the temperature of the cooling halo. Since a subset of the pure line emitters were heated by SN, there are also metals present in the system. For some galaxies we see faint O I, O II, O III, Ne III, and S II lines. The lowest ionization state lines tend to be the brightest of the metal lines and we highlight the notable lack of He II emission.

In most cases, these cooling haloes have $M_{UV} = 0$ (see the bottom panel of Figure 11), well beyond current capabilities. But the extreme tail of the UV magnitude distribution reaches -11 which is within reach of JWST when coupled with strong gravitational lensing.

4.2. Starbursting, Extreme Emission Line Galaxies

As the star formation rates within galaxies rise after the transition from Pop. III to Pop. II star formation due to the enhanced cooling from metals compared to H $_2$, a small fraction of galaxies undergo catastrophic bursts of star formation. Such bursts manifest spectroscopically as extreme emission line galaxies (EELGs). A stacked example of EELGs is shown as the red clump in the bottom right panel of Figure 8. While EELGs have been observed at all redshifts, for example blueberries at $z \sim 0$ ([Yang et al. 2017](#)), green peas at $z \sim 0.3$ ([Cardamone et al. 2009](#)), as well as others up to $z \sim 1$ ([Amorín et al. 2015](#)), they are certainly more common at high redshifts ([Boyett et al. 2024](#); [Davis et al. 2024](#); [Llerena et al. 2024](#)).

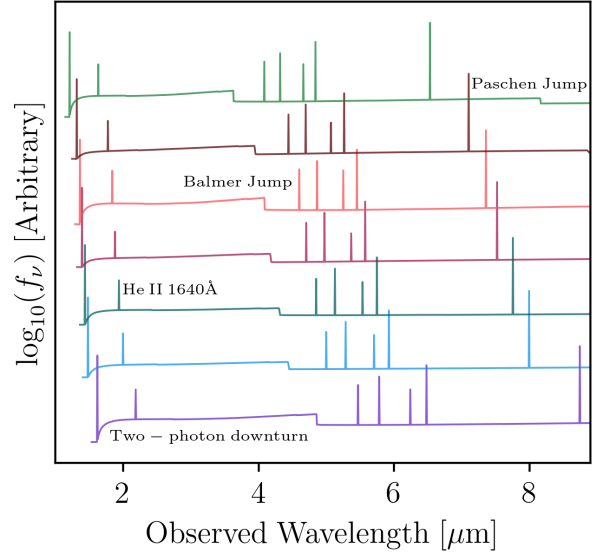


FIG. 10.— Seven example bright Pop. III galaxy spectra. Spectra are shown in f_ν and have been renormalized onto an arbitrary scale for clarity. The brightest Pop. III systems all show strong He II emission in the UV and optical, a Balmer (and Paschen) jump, and a downturn in the UV just redward of Ly α from two-photon emission.

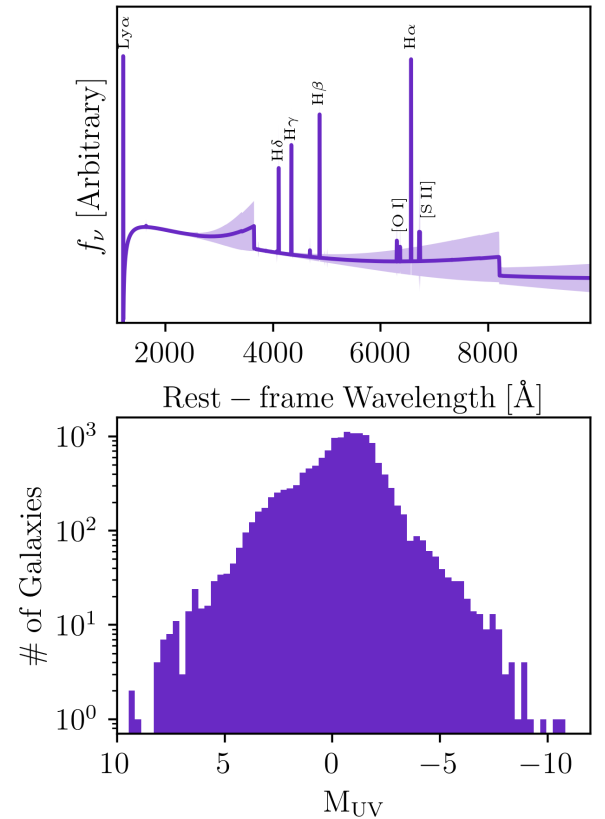


FIG. 11.— (Top) Median stacked spectra of cooling haloes (pure line emitters) across all simulations. The shaded region represents the 1σ standard deviation of the population. (Bottom) Histogram of the UV magnitudes of the cooling haloes.

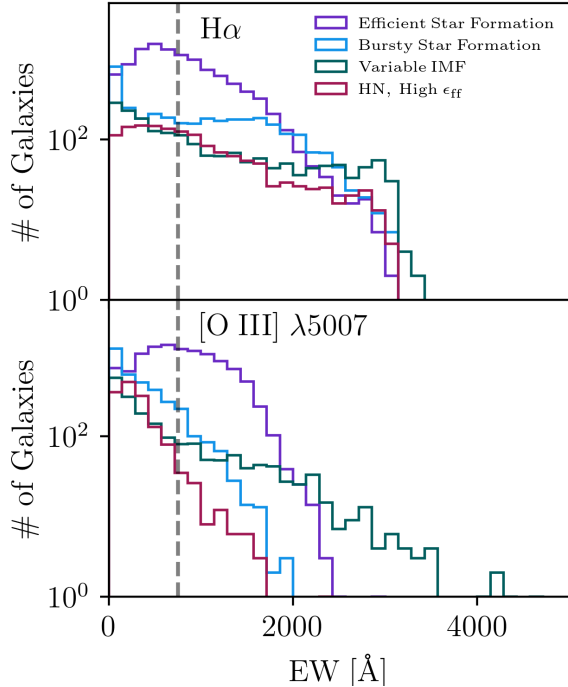


FIG. 12.— Histograms of $H\alpha$ (top) and $[O\ III]\ \lambda 5007$ (bottom) equivalent width for galaxies with UV magnitudes brighter than -15 . The dashed vertical line represents an EW of $750\ \text{\AA}$ which is our threshold for an EELG.

4.2.1. Emission Line Properties of EELGs

The MEGATRON simulations produce numerous EELGs, as shown in Figure 12, where we plot histograms of $H\alpha$ (top) and $[O\ III]\ \lambda 5007$ (bottom) equivalent widths for galaxies with UV magnitudes brighter than -15 . Following Boyett et al. (2024), we define an EELG as having an $H\alpha$ or $[O\ III]\ \lambda 5007$ EW $> 750\ \text{\AA}$. The fraction of bright galaxies that would be considered EELGs are similar across all runs at $\sim 50\%$ ⁹; however, the distributions of EWs are very different. For example, the efficient star formation $H\alpha$ distribution is peaked at close to $750\ \text{\AA}$ while the bursty star formation and variable IMF models have much flatter distributions. Because the efficient star formation simulation forms more stars overall compared to the other two models it produces the most EELGs.

A key result is that in the variable IMF simulation, $[O\ III]$ EWs can reach $> 4000\ \text{\AA}$. This occurs for galaxies in the magnitude range $\lesssim -19$ undergoing extreme bursts of star formation. Simulations historically struggle to reproduce the tail of high $[O\ III]$ EWs observed at high redshift (e.g., Ceverino et al. 2021; Wilkins et al. 2023). However, two physical effects differentiate the variable IMF model from the other simulations. First, HN in the variable IMF simulation inject a significant amount of oxygen which can increase the EW in scenarios where the oxygen emission is limited by metallicity¹⁰ and not by strong cooling. Second, the stellar popu-

⁹ Note that the EELG fraction is highly sensitive to the stellar mass distribution.

¹⁰ One may assume that the same physics would lead to higher $[O\ III]$ EWs in the HN, High ϵ_{ff} simulation; however, the stars form at much lower density in this run. Extreme densities are a requirement for forming the most extreme EELGs.

lations in the variable IMF simulation have higher mass loss which suppresses the stellar continuum at older ages. Both of these effects combined may help explain the high $[O\ III]$ EW galaxies seen in observations.

4.2.2. Star Formation Histories of EELGs

Analyzing the star formation histories of EELGs demonstrates how extreme they are. In the top panel of Figure 13 we show the median, normalized (by their most recent 1 Myr-averaged SFR), stacked star formation histories for galaxies with different $H\alpha$ EW for three of the high-redshift suite simulations. For this exercise, we only consider galaxies with $M_{UV} < -15$. While the EELG population with $H\alpha$ EW $> 1,800\ \text{\AA}$ sees a rapid rise in SFR where the current value is typically > 100 times more than what it was 20 Myr ago, an ordinary bright galaxy with an $H\alpha$ EW $\sim 400\ \text{\AA}$ sees almost no difference in SFR over the past 20 Myr. In contrast, the lowest $H\alpha$ EW galaxies have significant downturns in their recent star formation of up to a factor of 1,000 in the past 50 Myr in the most extreme examples. From this figure, we see how the $H\alpha$ EW is an extremely strong predictor for the typical SFH and sSFR of a galaxy (see also e.g., Fumagalli et al. 2012; Marmol-Queraltó et al. 2016; Khostovan et al. 2024). While the details of the SFH depend on the feedback and star formation model (compare the SFHs in the three panels of Figure 13), in all cases EELGs require an extreme recent burst of star formation.

4.2.3. Continuum Properties of EELGs

The differences in the emission line properties of EELGs are also reflected in the continuum features of the spectra. In the bottom panel of Figure 13 we show the normalized, stacked spectra for galaxies in bins of $H\alpha$ EW for the bursty star formation model¹¹. The most noticeable difference of the EELGs compared to lower EW galaxies (apart from the strength of the emission lines) occurs near H_∞ at $\sim 3645\ \text{\AA}$.

As discussed in Raiter, Schaerer & Fosbury (2010); Byler et al. (2017); Katz et al. (2024a), a unique property of EELGs, and more generally galaxies with very young stellar populations, is that the nebular continuum emission can contribute significantly in both the rest-frame UV and optical. In general, the nebular continuum causes the UV magnitude to increase while the spectra reddens (e.g., Bouwens et al. 2010; Dunlop et al. 2013). This is because the H I free-bound increases towards redder wavelengths, while the two-photon continuum peaks in the rest-frame UV. The nebular continuum is amplified when allowing the mass of the most massive star in the IMF to increase (Raiter, Schaerer & Fosbury 2010; Katz et al. 2024a; Schaerer et al. 2025).

The change in magnitude of the spectrum of each galaxy at $1500\ \text{\AA}$ and $3640\ \text{\AA}$ due to the addition of the nebular continuum as a function of intrinsic UV magnitude is shown in Figure 14. Focusing on a UV magnitude of ~ -18 , there is a wide diversity in nebular contributions. For the typical bright ($M_{UV} < -15$) galaxy in the efficient star formation and variable IMF models, the change in M_{UV} is ~ 0.05 magnitudes which would

¹¹ The other models are not noticeably different.

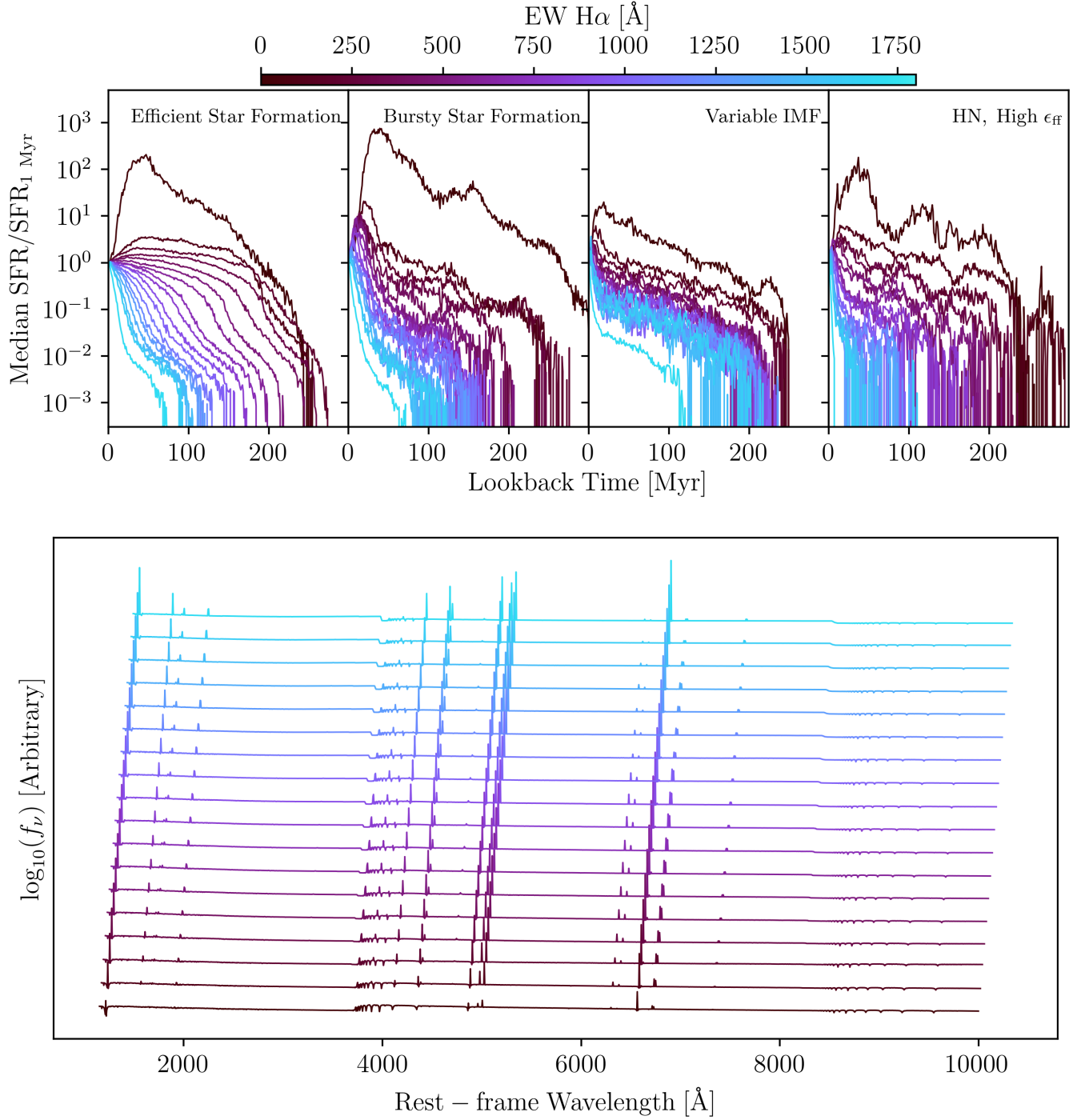


FIG. 13.— (Top) Median normalized star formation histories of bright galaxies with different H α EWs in each simulation. The lines are coloured by the minimum H α EW in the bin. (Bottom) Normalized stacked spectra for galaxies in each H α EW bin in the bursty star formation simulation. For clarity, we have offset each spectrum by 20 Å from the previous. In both panels, we consider only galaxies with $M_{\text{UV}} < -15$. Over a 50 Myr period, in the most extreme cases, the median stacked star formation histories can differ by a factor of $> 10^4$ between the different EW bins. Such drastic changes in SFR are reflected in the shape of the continuum and in the strength of the emission lines.

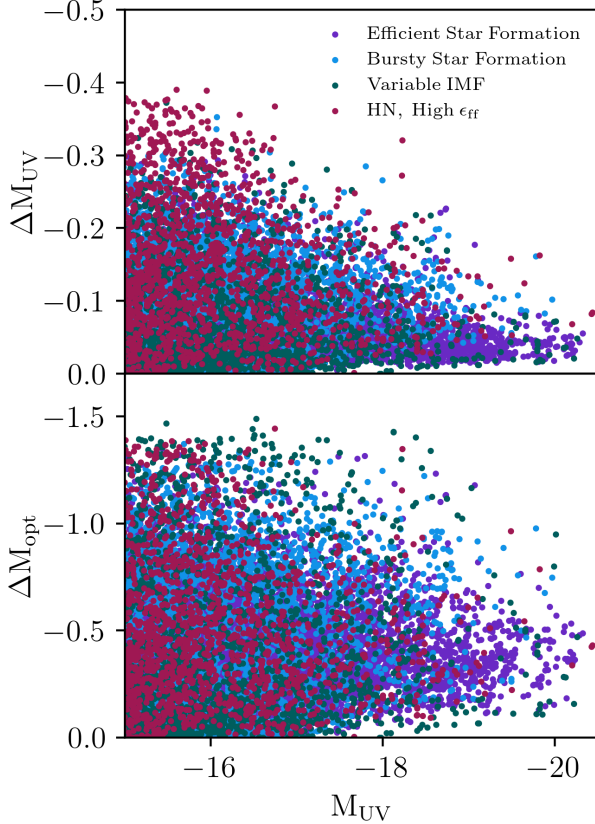


FIG. 14.— Contribution of the nebular continuum to the UV magnitude at 1500 Å (top) or the optical magnitude at 3640 Å (bottom) as a function of intrinsic UV magnitude. We consider all galaxies with $M_{UV} < -15$ at all redshifts.

have no impact on the UV luminosity. This increases to ~ 0.1 magnitudes for the other two models. However, for the galaxies undergoing the most extreme starbursts, the effect can be as high as 0.4 magnitudes for our assumed SSP models. The impact of the nebular continuum on the UV magnitudes of high-redshift galaxies therefore must not be ignored, especially for the most extreme systems which are typically the first to be detected.

The strength¹² and shape of the nebular contribution to the overall spectrum is primarily sensitive to three parameters: the specific star formation rate, the density of the ISM, and ξ_{ion} . The nebular continuum is strongest when ξ_{ion} is maximized (when there are more ionizing photons per unit 1500 Å luminosity), the LyC escape fraction is low (such that all ionizing photons are converted into nebular continuum), and the density is low (l -changing collisions shift electrons from the $2s$ to $2p$ state, reducing two-photon emission). In our models, ξ_{ion} is limited by our adopted SSP models; however, if high-redshift stars are hotter either due to lower metallicity and being more massive (Schaerer 2002), or due to peculiar chemical abundance patterns (Katz et al. 2024b), the nebular contribution, particularly in the UV, can become much more significant. Most stars form in embedded in dense clouds so LyC leakage rarely impacts

¹² Here we refer to both the overall normalization and with respect to the stellar continuum

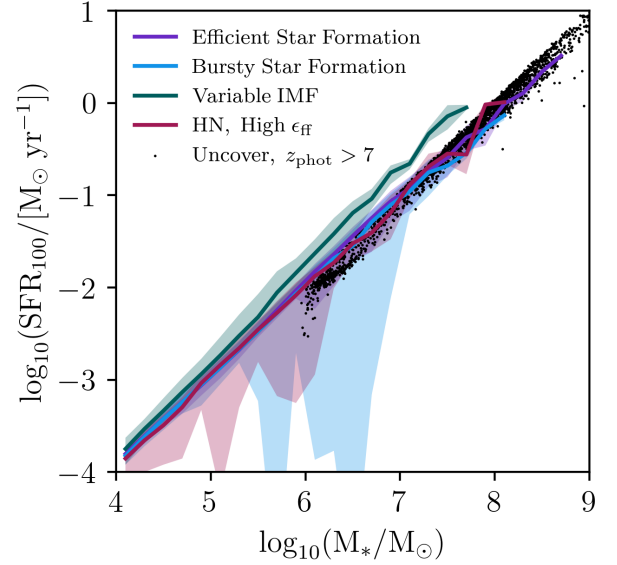


FIG. 15.— Star formation main-sequence averaged over 100 Myr for all galaxies in each high-redshift suite simulation. We show the median and 1σ scatter for each simulation. For comparison, we show SED fitting results from photometrically selected $z > 7$ galaxies from the UNCOVER survey.

the nebular continuum. However, the ISM in our models can reach densities of $\gtrsim 10^5 \text{ cm}^{-3}$, particularly in galaxies undergoing extreme bursts of star formation, which is above the critical density for H I two-photon emission. This latter effect leads to a wide diversity in spectral slope. If we consider the strongest EELGs with $H\alpha \text{ EW} > 1500 \text{ Å}$, we find a trend such that intrinsic UV slope scales with the mean gas density of a galaxy as $\beta = 0.038 \log_{10}(n_H/\text{cm}^{-3}) - 2.58$, i.e. galaxies with denser gas have redder slopes. Our simulations clearly demonstrate how the detailed properties of the ISM impact not only the emission lines but also the shape of the observed continuum.

4.3. Starforming Galaxies

The overwhelming majority (more than half) of galaxies in our simulations can be described as being starforming¹³. A stacked spectra of these galaxies is shown in brown in Figure 8. These galaxies have strong emission lines and blue UV slopes, and represent the less-extreme counterparts to the galaxies described in the previous section.

The starforming galaxy population exhibits less extreme scatter in their star formation rate as a function of stellar mass, which maintains a blue β , but lowers the EWs. Because of the less extreme changes in SFR, these galaxies form a tight main-sequence (e.g. Speagle et al. 2014; Schreiber et al. 2015). This main-sequence has now been characterized up to $z \sim 10$ with JWST data (e.g. Roberts-Borsani et al. 2024).

MEGATRON galaxies also exhibit a tight main-sequence, especially when averaged over long (~ 100 Myr) time scales. We show this main-sequence and its 1σ scatter for each high-redshift simulation in Figure 15. Nearly

¹³ This definition is qualitative and entirely based on the UMAP decomposition

independent of feedback model and star formation criteria, the galaxies all follow the same main sequence, which is in excellent agreement with SED fitting results from photometrically selected $z > 7$ galaxies from the UNCOVER survey (Wang et al. 2024; Weaver et al. 2024; Suess et al. 2024; Furtak et al. 2023). The exception to this trend is the variable IMF simulation where the enhanced mass loss causes an offset towards lower stellar mass. The SED fitting codes typically assume a fixed IMF and thus it is expected that the variable IMF simulation is offset.

Even though the median star formation main sequence are consistent across all runs with a fixed stellar IMF, the scatter in the relation depends substantially on the star formation and feedback model. For example, at stellar masses below $10^7 M_\odot$, the scatter becomes extremely large for the bursty star formation simulation. This is because the strong feedback can suppress star formation for time scales longer than 100 Myr. Hence the ratio of galaxies on and off the main-sequence is sensitive to the physics, even if the location of the main-sequence is not. The galaxies that have suppressed SFRs will be discussed in the next sections.

4.4. “Mini-Quenched” Galaxies

The star formation rates required to power an EELG are unsustainable for long periods of time. This is because these galaxies rapidly deplete their ISM gas reservoirs through star formation and intense, temporally coincident feedback can expel the residual ISM gas from the central parts of the galaxy. A sample of “Mini-Quenched” galaxies that have seen a relatively rapid and sustained drop in their SFR have been identified by JWST for having blue UV slopes, no emission lines, and Balmer absorption features (e.g., Looser et al. 2024; Trussler et al. 2025). The MEGATRON simulations contain a small population of such systems and their stacked spectra are shown in green in Figure 8.

Mini-quenched galaxies can be selected by having a low ratio of gas to stellar mass and a near-zero sSFR as shown in Figure 16. Low values of sSFR indicate that the galaxies are not currently undergoing any star formation and a low ratio of gas to stellar mass selects systems that have expelled all of their gas. Indeed, a population of galaxies emerges with extremely low sSFR (consistent with being 0) and very low gas fractions.

Two example mini-quenched galaxy spectra are shown in the top panel of Figure 17. These two galaxies have halo virial masses of $10^{8.6}$ and $10^9 M_\odot$, respectively, and gas masses that are $\lesssim 1\%$ of the cosmic baryon fraction. The low gas fractions and SFR result in a UV slope close to -2 (i.e. flat in f_ν). Neither nebular emission lines nor the continuum are visible in these spectra. The bottom panel of Figure 17 shows the star formation histories of the two galaxies where a sharp drop in the SFR can be seen 70–80 Myr ago. The extended period of no star formation is required to reach a UV slope close to -2 . What is noticeable about these galaxies is that in both, the SFR increased by a factor of ~ 1000 between 120 Myr ago and 80–90 Myr ago when the SFR reached its peak. Thus to form a mini-quenched galaxy in our simulations, we seem to require both a very rapid rise in star formation and a rapid decline due to stellar feedback expelling

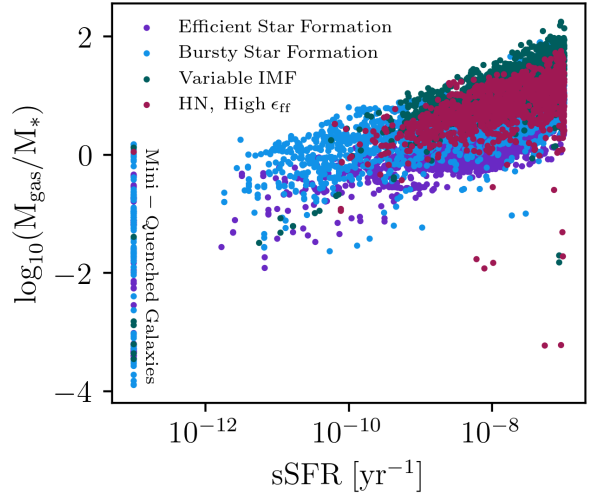


FIG. 16.— Ratio of gas to stellar mass as a function of sSFR for bright galaxies ($M_{UV} < -15$) in the high-redshift simulations. We have put a floor in sSFR at 10^{-13} yr^{-1} so that the quenched systems are visible in this diagram. To calculate sSFR, we use the 10 Myr-averaged SFR.

the gas.

One of the primary differences between the mini-quenched galaxies in the MEGATRON simulations and those observed by JWST is that the simulated analogs tend to be much fainter in the UV. In the simulations, these objects tend to have UV magnitudes in the range of -15 to -16 and stellar masses $\lesssim 10^{7.5} M_\odot$, consistent with results from the SERRA simulations (Gelli et al. 2025), and the gravitationally lensed object presented in Strait et al. (2023). In contrast, the object observed by Looser et al. (2024) is much more massive with an estimated stellar mass of $\sim 10^{8.5} M_\odot$.

Mini-quenched galaxies are extremely rare in our simulations. Because our simulations do not probe many massive objects at high redshift, it is possible that our sample size is not large enough to produce brighter mini-quenched objects. Gelli et al. (2024) speculated that SN feedback alone cannot quench the system in a way to reproduce the observed SED and posited that outflows from strong radiation pressure may help. Our simulations do include direct UV radiation pressure as well as multi-scattered radiation pressure on dust, but radiation pressure seems to play a minimal role. Dome et al. (2024) also found that other simulations (e.g., IllustrisTNG and VELA) fail to reproduce the SED of the Looser et al. (2024) galaxy, and thus forming massive mini-quenched galaxies seems to be a rather generic problem among simulations. Future simulations of more massive objects with our galaxy formation model may help elucidate the underlying physics driving this behavior.

4.5. Balmer Break and Post-Starburst Galaxies

The mini-quenched galaxies represent an extreme tail of the distribution of galaxies with aging stellar populations. Indeed they are clearly offset in our UMAP decomposition. A less extreme population of aging or post-starburst galaxies that have downturns in their SFR are identified by their Balmer break and weaker emission lines. A stacked spectra of these objects is shown in or-

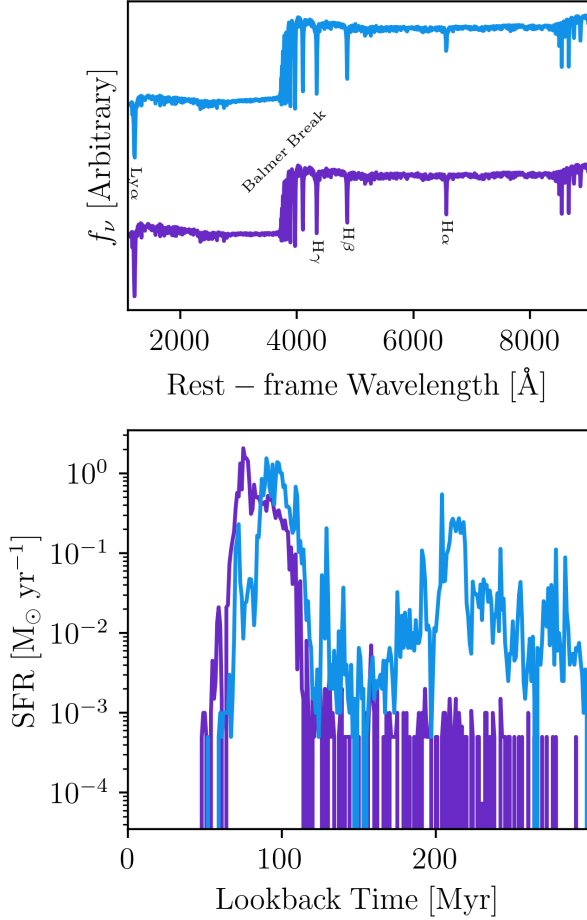


FIG. 17.— (Top) Spectra of two example mini-quenched galaxies. These spectra come from the efficient star formation and bursty star formation simulations. (Bottom) Star formation histories of the same two galaxies over the past 300 Myr.

ange in Figure 8. The UMAP decomposition identifies these objects as forming a smooth transition with the normal starforming galaxy population.

Even in the pre-JWST era, certain redshifts allowed for the photometric identification of the Balmer break (e.g., Hashimoto et al. 2018; Roberts-Borsani, Ellis & Laporte 2020). While some simulations can predict strong Balmer breaks at high redshift (e.g., Katz et al. 2019b; Wilkins et al. 2024), this is not the case for all (Binggeli et al. 2019).

We show the Balmer break strength measured as $F_{\nu,4200}/F_{\nu,3500}$ as a function of UV continuum luminosity for MEGATRON galaxies in Figure 18. For UV fluxes $\lesssim 10^{28} \text{ erg s}^{-1} \text{ Hz}^{-1}$, the MEGATRON simulations capture the diversity seen by JWST (Vikaeus et al. 2024). We do not find any very bright galaxies with strong Balmer breaks, likely due to the limited volume of the simulation. The brightest objects in our volume are all galaxies that deviate high on the star-formation main-sequence and have spectra with steep UV slopes and Balmer jumps.

Some of the Balmer break galaxies observed by JWST also exhibit strong emission lines (Kuruvanthodi et al. 2024). At lower redshifts, similar features are seen in

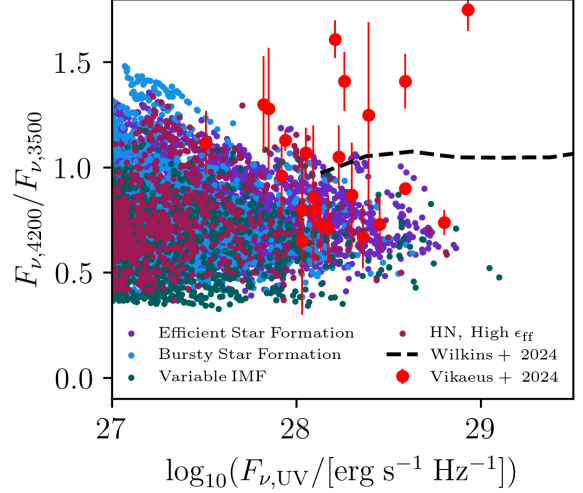


FIG. 18.— Balmer break strength as a function of UV continuum luminosity. For comparison, we show the sample of Balmer breaks measured for high-redshift JWST galaxies from Vikaeus et al. (2024) as well as the $z = 9$ unattenuated predictions from the FLARES simulations (Wilkins et al. 2024).

shocked post-starburst galaxies (Alatalo et al. 2016). Shocks are not always well resolved in our simulations due to the finite spatial resolution, and we do not explicitly post-process the simulation to account for unresolved shocks (see e.g., Hirschmann et al. 2023); however, we do find a subset of galaxies that exhibit Balmer breaks with post-starburst features and shock-like emission line ratios.

One way to search for post-starburst galaxies is to identify systems with strong low ionization state emission lines and UV slopes indicative of aging stellar populations. In Figure 19 we show UV slope as a function of $[\text{S II}] \lambda\lambda 6716, 6731/\text{H}\alpha$ for galaxies with UV magnitudes < -15 . Particularly in the efficient star formation simulation, there is a small population of galaxies that has $\beta \gtrsim -2$ and $[\text{S II}] \lambda\lambda 6716, 6731/\text{H}\alpha \gtrsim 0.5$. Such a high $[\text{S II}] \lambda\lambda 6716, 6731/\text{H}\alpha$ ratio is indicative of diffuse gas emission and/or shocks, while a UV slope close to -2 , in the absence of dust, indicates an aging stellar population.

The spectrum of the galaxy with the highest $[\text{S II}] \lambda\lambda 6716, 6731/\text{H}\alpha$ ratio and the reddest UV slope is shown in the top panel of Figure 20. There are indeed very strong $[\text{S II}]$ and $[\text{O I}]$ lines as well as $[\text{N II}]$ and $[\text{O II}]$, and a clear Balmer break. The SFH of this galaxy is very peculiar (as shown in the bottom panel). The system underwent two extreme bursts of star formation ~ 160 and 220 Myr ago reaching SFRs ~ 200 times the current value. Subsequently, the SFR has had a relatively steady decline. The old stellar populations are necessary for driving the Balmer break while the current star formation is key for producing the small amount of $[\text{O III}]$ emission seen in the galaxy. The halo hosting this object is relatively massive, with a virial mass nearing $10^9 M_\odot$ and we find no other objects with such peculiar spectra in any of the other simulations.

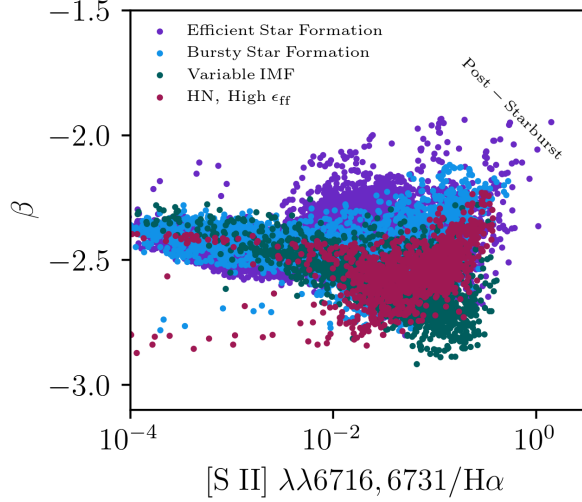


FIG. 19.— UV slope (β) as a function of $[\text{S II}]/\text{H}\alpha$ for galaxies with UV magnitudes brighter than -15 . Galaxies with the reddest (greatest) slopes and the highest $[\text{S II}]/\text{H}\alpha$ ratio represent a post-starburst population. This region is labelled qualitatively.

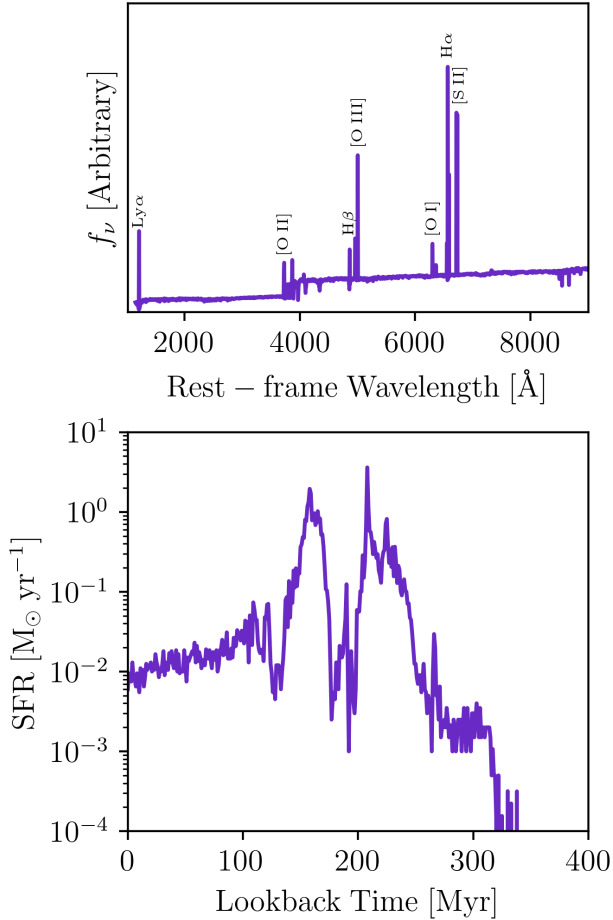


FIG. 20.— Example spectrum (top) and star formation history (bottom) of a post-starburst galaxy in the bursty star formation simulation.

5. DISCUSSION

We have presented an overview of the MEGATRON suite of cosmological radiation hydrodynamics simulations that focus on the Lagrange region around a Milky Way-mass environment at high redshift. This suite of simulations is the first to couple a detailed non-equilibrium thermochemistry network of primordial species, metals, and molecules with high enough resolution to predict the intrinsic spectra of galaxies including their stellar and nebular emission.

Direct comparisons with observations remain the primary tool for constraining the physics of galaxy formation. For example, a key milestone of numerical simulations was the ability to reproduce the Hubble sequence of galaxy morphology at $z = 0$ (e.g., [Vogelsberger et al. 2014](#); [Schaye et al. 2015](#); [Dubois et al. 2016](#)) and these simulations were remarkably successful in reproducing numerous other observed or inferred characteristics of galaxies such as the stellar mass function. However, historically at high redshift, besides the UV luminosity function few observational constraints existed on the detailed properties of galaxies in the epoch of reionization. This effort was primarily driven by ALMA (e.g., [Pentericci et al. 2016](#); [Bradać et al. 2017](#); [Hashimoto et al. 2019](#); [Harikane et al. 2020](#); [Carniani et al. 2020](#)) or the occasional UV line detection from large ground-based facilities (e.g., [Stark et al. 2017](#); [Mainali et al. 2017](#)). Such observations were time-consuming for very small samples of galaxies.

The successful launch of JWST has revolutionized the field of high-redshift astronomy from an observational perspective. This presents an opportunity to test predictions from simulations in a previously unexplored manner and represents new data to calibrate underlying physical models of simulations. One of the key technological advantages of JWST is its spectral capabilities that probe the properties of the high-redshift ISM. Just like the morphological diversity presents a strong constraint on the physics of galaxy formation, so does spectral diversity. However, few large scale simulations even attempt to resolve the ISM (e.g., [Vogelsberger et al. 2014](#); [Schaye et al. 2015](#); [Dubois et al. 2016](#); [Pillepich et al. 2018](#); [Lovell et al. 2021](#)) and those that do (e.g., [O’Shea et al. 2015](#); [Ceverino, Glover & Klessen 2017](#); [Rosdahl et al. 2018](#); [Ma et al. 2018](#); [Trebitsch et al. 2021](#); [PalloTTini et al. 2022](#); [Kannan et al. 2025](#)) neglect the physics that gives rise to all of the strong emission lines seen in JWST spectra. For these reasons, recent high-resolution simulations use detailed and computationally expensive post-processing to model galaxy spectra. As discussed in [Katz et al. \(2023b\)](#), numerous different, but physically motivated assumptions are often made in post-processing that can limit the fidelity of predicted spectra. Hence, attempting to predict the diversity of galaxy spectra seen in the early Universe by JWST under non-equilibrium conditions, represents one of the primary scientific motivations for the MEGATRON project. As shown in [Figure 8](#), much of the spectral diversity results from the complex star formation histories driven by gas accretion, mergers, dynamical instabilities, and feedback processes present at high redshift. This qualitative agreement between theory and observation represents only a first step for constraining models of galaxy formation. A more

detailed comparison on spectral diagnostics will be presented in [Choustikov et al. *in prep.*](#) where we expand on the results from [Katz et al. \(2024c\)](#) and demonstrate how observed spectra can be used as a key test of subgrid models.

The results from our simulations are subject to numerous caveats. We refrain from discussing the limitations of numerical methods, in particular those related to finite spatial and mass resolution, subgrid models, and unresolved Strömgren spheres, as these are discussed at length in the context of the MEGATRON model in [Katz et al. \(2024c\)](#). Likewise, we acknowledge the exclusion of magnetic fields and cosmic rays, which represent an important pressure term and heating/ionization source, respectively in the ISM. These will be addressed in upcoming work. Similarly, we do not attempt to model massive black holes, primarily due to the uncertainties in their seeding/formation channels, how they accrete, their feedback, and their SEDs, despite their unexpectedly high abundance and mass ratios compared to the stellar content of galaxies at high redshift (e.g., [Maiolino et al. 2024; Matthee et al. 2024b](#)). Rather, we highlight uncertainties in our modeling that can have the largest impacts on the observed spectral properties of galaxies.

First, stellar SEDs, particularly in the ionizing regime are highly uncertain at low metallicity. For example, switching SEDs from single to binary stars and drastically change the reionization history ([Ma et al. 2016; Rosdahl et al. 2018](#)). Within the context of our modeling, excess ionizing photons or a change in spectral hardness impacts both the EW and ratios of strong emission lines as well as the contribution of the nebular continuum to the total SED. The stellar IMF is also highly uncertain at high redshift and while we varied the upper mass slope in the variable IMF model, we have not considered significant changes to the upper mass limit. Atmospheric models for such massive, metal poor stars are scant (see e.g., [Martins et al. 2020](#)) and we are thus limited by their availability. However, in the supermassive regime, the luminosities can reach $\sim 10^9 L_\odot$ and can dominate the SEDs of the lower-mass galaxies at high redshift. Furthermore, we have not considered deviations from solar abundance patterns, despite metal abundances of the star particles being highly non-solar. α -enhancements have been considered in population synthesis models (e.g., [Byrne et al. 2022; Park et al. 2024](#)) and recently [Katz et al. \(2024b\)](#) discussed the importance of He enhancements for high-redshift star clusters, but none of these effects are considered in this work. Moreover, X-ray binaries are known to be more important at low metallicity (e.g. [Mapelli et al. 2010; Kaaret, Schmitt & Gorski 2011; Basu-Zych et al. 2013; Douna et al. 2015; Saxena et al. 2021](#)) which would be very relevant for the systems simulated here, and certain JWST observations are better explained when the effects of X-ray binaries are accounted for ([Katz et al. 2023c](#)). As a final point, we do not model the spectra of Pop. II stars individually, but rather we assume an average SED. In reality, especially in the dwarf galaxy regime, there may be local fluctuations in the SED properties due to IMF sampling that may be important ([Stanway & Eldridge 2023](#)) and are not captured here.

Second, large uncertainties remain in the chemical yields ([Buck et al. 2021](#)) and their depletion onto dust.

These uncertainties are particularly important in the context of our work because gas cooling is highly sensitive to presence of individual species and their depletion rates, as are the observed metal emission lines. Cooling not only impacts the efficiency by which gas can collapse and form stars or how fast shocks instigated by SN can cool, but most of the strong metal emission lines have emissivities that are exponentially sensitive to electron temperature. Small changes in the cooling can have an outsized impact on the predicted emission. While we are confident that for a given set of abundances, our cooling model is accurate¹⁴, if the abundance patterns that our adopted yields predict are not consistent with real galaxies, we may be unable to reproduce some of the spectral characteristics at high redshift.

Third, little is known about Pop. III stars because they have never been directly observed. We have adopted a plausible model that has been used previously ([Wise et al. 2012; Kimm et al. 2017](#)); however, without observational guidance, little progress can be made on their detailed properties. This means that our predictions on the observable properties of the Pop. III-Pop. II transition may be highly uncertain. If JWST does not provide more constraints on Pop. III stars, the ELT may present as a promising alternative ([Grisdale et al. 2021](#)).

6. CONCLUSIONS

With the caveats discussed in the previous section in mind, we highlight our primary conclusions.

1. Much of the spectral diversity seen by JWST naturally emerges within the context of a Λ CDM cosmology and reasonable choices for subgrid models (that in our case have been benchmarked to $z = 0$). This is despite the fact that our initial conditions represent a highly biased environment and result from highly varied and complex star formation histories. Nevertheless, significant work remains, both observationally to obtain spectroscopic samples of high-redshift galaxies with a well defined selection function, and the theoretically to not only reproduce the diversity, but also the proportions of each class of spectra. Furthermore, our simulations still struggle to produce certain galaxies like massive mini-quenched systems and other peculiar objects including the high-redshift nitrogen emitters (e.g. [Bunker et al. 2023; Castellano et al. 2024; Naidu et al. 2025](#)) and Little Red Dots ([Matthee et al. 2024a](#)).
2. While individual haloes rapidly transition from Pop. III to Pop. II star formation, across the entire Lagrange region, Pop. III star formation continues to the lowest redshift ($z = 8.5$) currently reached by our simulations. The Pop. III spectra tend to be redder than those of Pop. II galaxies due to the strong nebular contribution and high residual ISM gas densities. Most Pop. III galaxies are too UV faint to be observed, even with optimistic assumptions about strong gravitational lensing. We identify a (new) class of haloes without any stars on the

¹⁴ Or at the very least is able to reproduce more detailed photoionization codes such as CLOUDY.

main-sequence where primordial and metal cooling radiation dominate their spectra. The brightest examples of such galaxies appear immediately after a Pop. III star formation event.

3. The most extreme emission line galaxies form when the instantaneous star formation rates of galaxies spike to values $\gtrsim 100\times$ what they were averaged over the previous 50 Myr. $\sim 50\%$ of the simulated galaxies fit the observational definition of being an EELG, which is qualitatively consistent with observations that demonstrate their increased occurrence towards high redshift. Most of our simulations struggle to reproduce the highest [O III] EWs seen in observations; however, we find that the inclusion of hypernova and a top-heavy IMF may help remedy this historic disagreement between simulations and observations.
4. Independent of star formation or feedback model, our simulations easily reproduce the observed star formation main-sequence. This is due to the fact that the majority of the low-mass galaxies in both the simulations and observations seemed to have assembled most of their stellar mass within a 100 Myr time period. Therefore the scatter in this relation and comparing with SFR on shorter timescales will provide a stronger constraint on the physics of high-redshift galaxy formation.
5. Our simulations struggle to produce bright galaxies with strong Balmer breaks and massive mini-quenched galaxies. Part of this discrepancy may be due to the limited volume of the simulations and thus our feedback model and subgrid physics requires further testing in more extreme environments. Such spectra are much more easily reproduced for faint galaxies that exhibit more stochastic star formation histories.

Here we have focused primarily on high-redshift galaxy formation with the MEGATRON simulations. Hence we have only addressed two of the four scientific goals of the project. The connection with low redshift via near-field cosmology and the CGM towards Cosmic Noon are explored in companion papers (Rey et al. 2025; Cadiou et al. 2025). Furthermore, data products from the MEGATRON simulations will be released to the wider community in the hope that it aids in both interpreting observations and helps constrain the key physics of early galaxy formation.

ACKNOWLEDGMENTS

This work made extensive use of the dp265, dp016, dp373, and dp379 projects on the DiRAC ecosystem. HK is particularly grateful to Christopher Mounford and Alastair Basden for support on DIaL3 and Cosma8, respectively. HK and the MEGATRON team are especially thankful for the support on Glamdring provided by Jonathan Patterson. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The

equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. This work also used the DiRAC Data Intensive service at Leicester, operated by the University of Leicester IT Services, which forms part of the STFC DiRAC HPC Facility. The equipment was funded by BEIS capital funding via STFC capital grants ST/K000373/1 and ST/R002363/1 and STFC DiRAC Operations grant ST/R001014/1. This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk). DiRAC is part of the National e-Infrastructure. This work has made use of the Infinity Cluster hosted by Institut d’Astrophysique de Paris. We thank Stephane Rouberol for running smoothly this cluster for us.

The material in this manuscript is based upon work supported by NASA under award No. 80NSSC25K7009. HK acknowledges support from FACCTS. AS and AJC acknowledge funding from the “FirstGalaxies” Advanced Grant from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No.789056). TK is supported by the National Research Foundation of Korea (RS-2022-NR070872 and RS-2025-00516961) and by the Yonsei Fellowship, funded by Lee Youn Jae. AS acknowledges support from the Science and Technology Facilities Council (STFC) for a PhD studentship. FRM is supported by the Kavli Institute for Cosmological physics at the University of Chicago through an endowment from the Kavli Foundation and its founder Fred Kavli. GCJ acknowledges support by the Science and Technology Facilities Council (STFC), by the ERC through Advanced Grant 695671 “QUENCH”, and by the UKRI Frontier Research grant RISEandFALL. KM acknowledges the Flemish Fund for Scientific Research (FWO-Vlaanderen), Grant number 1169822N. M.S. acknowledges the support from the Swiss National Science Foundation under Grant No. P500PT_214488. NC acknowledges support from the Science and Technology Facilities Council (STFC) for a PhD studentship. OA acknowledges support from the Knut and Alice Wallenberg Foundation, the Swedish Research Council (grant 2019-04659), the Swedish National Space Agency (SNSA Dnr 2023-00164), and the LMK foundation. The authors also acknowledge financial support from Oriel College’s Research Fund.

The authors thank Romain Teyssier and Leo Michel-Dansac for both developing and open-sourcing RAMSES and RASCAS, respectively. We thank the developers and maintainers of PYNBODY (Pontzen et al. 2013, 2023), NUMPY (van der Walt, Colbert & Varoquaux 2011; Harris et al. 2020), SCIPY (Virtanen et al. 2020), JUPYTER (Ragan-Kelley et al. 2014), MATPLOTLIB (Hunter 2007), the Astrophysics Data Service and the arXiv pre-print repository for providing open-source software and services that were used extensively in this work.

REFERENCES

- Agertz O., Kravtsov A. V., 2015, *ApJ*, 804, 18
- Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, *ApJ*, 770, 25
- Agertz O. et al., 2021, *MNRAS*, 503, 5826
- Alatalo K. et al., 2016, *ApJS*, 224, 38
- Aldrovandi S. M. V., Pequignot D., 1973, *A&A*, 25, 137
- Amorín R. et al., 2015, *A&A*, 578, A105
- Arata S., Yajima H., Nagamine K., Abe M., Khochfar S., 2020, *MNRAS*, 498, 5541
- Arnaud M., Raymond J., 1992, *ApJ*, 398, 394
- Arnaud M., Rothenflug R., 1985, *A&AS*, 60, 425
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Baczynski C., Glover S. C. O., Klessen R. S., 2015, *MNRAS*, 454, 380
- Badnell N. R., 2006, *ApJS*, 167, 334
- Badnell N. R. et al., 2003, *A&A*, 406, 1151
- Bakes E. L. O., Tielens A. G. G. M., 1994, *ApJ*, 427, 822
- Barragán P., Errea L. F., Méndez L., Rabadán I., Riera A., 2006, *ApJ*, 636, 544
- Barrow K. S. S., Wise J. H., Norman M. L., O'Shea B. W., Xu H., 2017, *MNRAS*, 469, 4863
- Basu-Zych A. R. et al., 2013, *ApJ*, 774, 152
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
- Bhagwat A., Napolitano L., Pentericci L., Ciardi B., Costa T., 2024, *arXiv e-prints*, arXiv:2408.16063
- Bialy S., Sternberg A., 2019, *ApJ*, 881, 160
- Binggeli C. et al., 2019, *MNRAS*, 489, 3827
- Black J. H., 1981, *MNRAS*, 197, 553
- Black J. H., Dalgarno A., 1977, *ApJS*, 34, 405
- Bouwens R. J. et al., 2010, *ApJ*, 708, L69
- Boyett K. et al., 2024, *MNRAS*, 535, 1796
- Bradač M. et al., 2017, *ApJ*, 836, L2
- Brauer K. et al., 2025, *ApJ*, 980, 41
- Bromm V., Ferrara A., Coppi P. S., Larson R. B., 2001, *MNRAS*, 328, 969
- Buck T., Rybizki J., Buder S., Obreja A., Macciò A. V., Pfrommer C., Steinmetz M., Ness M., 2021, *MNRAS*, 508, 3365
- Bunker A. J. et al., 2024, *A&A*, 690, A288
- Bunker A. J. et al., 2023, *A&A*, 677, A88
- Burton M. G., Hollenbach D. J., Tielens A. G. G. M., 1990, *ApJ*, 365, 620
- Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, 840, 44
- Byrne C. M., Stanway E. R., Eldridge J. J., McSwiney L., Townsend O. T., 2022, *MNRAS*, 512, 5329
- Cadiou C., Dubois Y., Pichon C., 2019, *Astronomy & Astrophysics*, 621, A96
- Cadiou C., Katz H., Rey M., the MEGATRON team, 2025, *arXiv e-prints*
- Cameron A. J., Katz H., Rey M. P., Saxena A., 2023a, *MNRAS*, 523, 3516
- Cameron A. J., Katz H., Witten C., Saxena A., Laporte N., Bunker A. J., 2023b, *arXiv e-prints*, arXiv:2311.02051
- Cameron A. J., Katz H., Witten C., Saxena A., Laporte N., Bunker A. J., 2024, *MNRAS*, 534, 523
- Cameron A. J. et al., 2023c, *A&A*, 677, A115
- Cardamone C. et al., 2009, *MNRAS*, 399, 1191
- Carniani S. et al., 2020, *MNRAS*, 499, 5136
- Carniani S. et al., 2024, *Nature*, 633, 318
- Castellano M. et al., 2025, *arXiv e-prints*, arXiv:2504.05893
- Castellano M. et al., 2024, *ApJ*, 972, 143
- Cen R., 1992, *ApJS*, 78, 341
- Ceverino D., Glover S. C. O., Klessen R. S., 2017, *MNRAS*, 470, 2791
- Ceverino D., Hirschmann M., Klessen R. S., Glover S. C. O., Charlot S., Feltre A., 2021, *MNRAS*, 504, 4472
- Chan T. K., Richings A. J., Theuns T., Liu Y., Schaller M., Ivkovic M., 2025, *arXiv e-prints*, arXiv:2508.13277
- Chemerynska I. et al., 2025, *arXiv e-prints*, arXiv:2509.24881
- Chemerynska I. et al., 2023, *arXiv e-prints*, arXiv:2312.05030
- Clark P. C., Glover S. C. O., Klessen R. S., Bromm V., 2011, *ApJ*, 727, 110
- Commerçon B., Debout V., Teyssier R., 2014, *A&A*, 563, A11
- Cullen F. et al., 2025, *arXiv e-prints*, arXiv:2501.11099
- Cullen F. et al., 2024, *MNRAS*, 531, 997
- Davis K. et al., 2024, *ApJ*, 974, 42
- de Graaff A. et al., 2024, *arXiv e-prints*, arXiv:2404.05683
- Dere K. P., Del Zanna G., Young P. R., Landi E., Sutherland R. S., 2019, *ApJS*, 241, 22
- D'Eugenio F. et al., 2025, *ApJS*, 277, 4
- Dijkstra M., 2009, *ApJ*, 690, 82
- Dome T., Tacchella S., Fialkov A., Ceverino D., Dekel A., Ginzburg O., Lapiner S., Looser T. J., 2024, *MNRAS*, 527, 2139
- Dopita M. A., Kewley L. J., Heisler C. A., Sutherland R. S., 2000, *ApJ*, 542, 224
- Douna V. M., Pellizza L. J., Mirabel I. F., Pedrosa S. E., 2015, *A&A*, 579, A44
- Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*
- Draine B. T., Bertoldi F., 1996, *ApJ*, 468, 269
- Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, 463, 3948
- Dunlop J. S. et al., 2013, *MNRAS*, 432, 3520
- Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *PASA*, 34, e058
- Federrath C., Klessen R. S., 2012, *ApJ*, 761, 156
- Ferland G. J. et al., 2017, *Rev. Mexicana Astron. Astrofis.*, 53, 385
- Field G. B., Goldsmith D. W., Habing H. J., 1969, *ApJ*, 155, L149
- Finkelstein S. L. et al., 2023, *arXiv e-prints*, arXiv:2311.04279
- Fujimoto S. et al., 2025, *arXiv e-prints*, arXiv:2501.11678
- Fumagalli M. et al., 2012, *ApJ*, 757, L22
- Furtak L. J. et al., 2023, *MNRAS*, 523, 4568
- Gardner J. P. et al., 2023, *PASP*, 135, 068001
- Gardner J. P. et al., 2006, *Space Sci. Rev.*, 123, 485
- Garg P., Narayanan D., Sanders R. L., Davé R., Popping G., Shapley A. E., Stark D. P., Trump J. R., 2024, *ApJ*, 972, 113
- Gelli V., Pallottini A., Salvadori S., Ferrara A., Mason C., Carniani S., Ginolfi M., 2025, *arXiv e-prints*, arXiv:2501.16418
- Gelli V., Salvadori S., Ferrara A., Pallottini A., 2024, *ApJ*, 964, 76
- Glazebrook K. et al., 2024, *Nature*, 628, 277
- Glover S. C. O., Abel T., 2008, *MNRAS*, 388, 1627
- Glover S. C. O., Clark P. C., 2012, *MNRAS*, 421, 116
- Glover S. C. O., Federrath C., Mac Low M. M., Klessen R. S., 2010, *MNRAS*, 404, 2
- Gnat O., Sternberg A., 2007, *ApJS*, 168, 213
- Gnedin N. Y., Tassis K., Kravtsov A. V., 2009, *ApJ*, 697, 55
- Gräfener G., Vink J. S., 2015, *A&A*, 578, L2
- Gray W. J., Scannapieco E., 2017, *ApJ*, 849, 132
- Grevesse N., Asplund M., Sauval A. J., Scott P., 2010, *Ap&SS*, 328, 179
- Grisdale K., Thatte N., Devriendt J., Pereira-Santaella M., Slyz A., Kimm T., Dubois Y., Yi S. K., 2021, *MNRAS*, 501, 5517
- Guillet T., Teyssier R., 2011, *Journal of Computational Physics*, 230, 4756
- Haiman Z., Thoul A. A., Loeb A., 1996, *ApJ*, 464, 523
- Harikane Y., Nakajima K., Ouchi M., Umeda H., Isobe Y., Ono Y., Xu Y., Zhang Y., 2024, *ApJ*, 960, 56
- Harikane Y. et al., 2020, *ApJ*, 896, 93
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Hashimoto T. et al., 2019, *PASJ*, 71, 71
- Hashimoto T. et al., 2018, *Nature*, 557, 392
- Heays A. N., Bosman A. D., van Dishoeck E. F., 2017, *A&A*, 602, A105
- Heger A., Woosley S. E., 2002, *ApJ*, 567, 532
- Hirano S., Hosokawa T., Yoshida N., Umeda H., Omukai K., Chiaki G., Yorke H. W., 2014, *ApJ*, 781, 60
- Hirschmann M. et al., 2023, *MNRAS*, 526, 3610
- Hirschmann M., Charlot S., Somerville R. S., 2023, *MNRAS*, 526, 3504
- Hollenbach D., McKee C. F., 1989, *ApJ*, 342, 306
- Hu W. et al., 2024, *ApJ*, 971, 21
- Hui L., Gnedin N. Y., 1997, *MNRAS*, 292, 27
- Hunter J. D., 2007, *Computing in Science & Engineering*, 9, 90
- Isobe Y. et al., 2023, *ApJ*, 959, 100
- Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., Schaerer D., Thuan T. X., 2021, *A&A*, 646, A138
- Izotov Y. I., Guseva N. G., Thuan T. X., 2011, *ApJ*, 728, 161
- Jaacks J., Thompson R., Finkelstein S. L., Bromm V., 2018, *MNRAS*, 475, 4396
- Jappes A. K., Glover S. C. O., Klessen R. S., Mac Low M. M., 2007, *ApJ*, 660, 1332
- Jin Y., Kewley L. J., Sutherland R. S., 2022, *ApJ*, 934, L8
- Joshi G. D., Pontzen A., Agertz O., Rey M. P., Read J., Pillepich A., 2025, *MNRAS*, 537, 3792
- Kaaret P., Schmitt J., Gorski M., 2011, *ApJ*, 741, 10
- 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*, arXiv:2502.20437
- Katz H., 2022, *MNRAS*, 512, 348
- Katz H. et al., 2024a, *arXiv e-prints*, arXiv:2408.03189
- Katz H. et al., 2019a, *MNRAS*, 487, 5902
- Katz H., Ji A. P., Telford G., Senchyna P., 2024b, *The Open Journal of Astrophysics*, 7, 106

- Katz H., Kimm T., Ellis R. S., Devriendt J., Slyz A., 2023a, *MNRAS*, 524, 351
- Katz H., Laporte N., Ellis R. S., Devriendt J., Slyz A., 2019b, *MNRAS*, 484, 4054
- Katz H. et al., 2022, arXiv e-prints, arXiv:2211.04626
- Katz H., Rey M. P., Cadiou C., Kimm T., Agertz O., 2024c, arXiv e-prints, arXiv:2411.07282
- Katz H. et al., 2023b, *The Open Journal of Astrophysics*, 6, 44
- Katz H. et al., 2023c, *MNRAS*, 518, 592
- Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, *ARA&A*, 57, 511
- Khostovan A. A., Malhotra S., Rhoads J. E., Sobral D., Harish S., Tilvi V., Coughlin A., Rezaee S., 2024, *MNRAS*, 535, 2903
- Kim C.-G., Ostriker E. C., 2015, *ApJ*, 802, 99
- Kim J.-G., Gong M., Kim C.-G., Ostriker E. C., 2023, *ApJS*, 264, 10
- Kimm T., Katz H., Haehnelt M., Rosdahl J., Devriendt J., Slyz A., 2017, *MNRAS*, 466, 4826
- Kingdon J. B., Ferland G. J., 1996, *ApJS*, 106, 205
- Kobayashi C., Umeda H., Nomoto K., Tominaga N., Ohkubo T., 2006, *ApJ*, 653, 1145
- Kokorev V. et al., 2024, arXiv e-prints, arXiv:2411.13640
- Koyama H., Inutsuka S.-I., 2000, *ApJ*, 532, 980
- Kroupa P., 2001, *MNRAS*, 322, 231
- Kumar S., Chen H.-W., 2025
- Kuruvanthodi A., Schaerer D., Marques-Chaves R., Korber D., Weibel A., Oesch P. A., Roberts-Borsani G., 2024, *A&A*, 691, A310
- Larkin M. M., Gerasimov R., Burgasser A. J., 2023, *AJ*, 165, 2
- Laseter I. H. et al., 2024, *A&A*, 681, A70
- Leitherer C. et al., 1999, *ApJS*, 123, 3
- Leung G. C. K. et al., 2023, *ApJ*, 954, L46
- Levermore C. D., 1984, *J. Quant. Spec. Radiat. Transf.*, 31, 149
- Limongi M., Chieffi A., 2018, *ApJS*, 237, 13
- Llerena M. et al., 2024, *A&A*, 691, A59
- Looser T. J. et al., 2024, *Nature*, 629, 53
- Lovell C. C., Vijayan A. P., Thomas P. A., Wilkins S. M., Barnes D. J., Irodotou D., Roper W., 2021, *MNRAS*, 500, 2127
- Lupi A., Pallottini A., Ferrara A., Bovino S., Carniani S., Vallini L., 2020, *MNRAS*, 496, 5160
- Luridiana V., Morisset C., Shaw R. A., 2015, *A&A*, 573, A42
- Ma X. et al., 2018, *MNRAS*, 478, 1694
- Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016, *MNRAS*, 459, 3614
- Mainali R., Kollmeier J. A., Stark D. P., Simcoe R. A., Walth G., Newman A. B., Miller D. R., 2017, *ApJ*, 836, L14
- Maiolino R. et al., 2024, *A&A*, 691, A145
- Maoz D., Graur O., 2017, *ApJ*, 848, 25
- Mapelli M., Ripamonti E., Zampieri L., Colpi M., Bressan A., 2010, *MNRAS*, 408, 234
- Marks M., Kroupa P., Dabringhausen J., Pawlowski M. S., 2012, *MNRAS*, 422, 2246
- Mármol-Queraltó E., McLure R. J., Cullen F., Dunlop J. S., Fontana A., McLeod D. J., 2016, *MNRAS*, 460, 3587
- Martins F., Schaerer D., Haemmerlé L., Charbonnel C., 2020, *A&A*, 633, A9
- Matthee J. et al., 2024a, *ApJ*, 963, 129
- Matthee J. et al., 2024b, arXiv e-prints, arXiv:2412.02846
- McClymont W. et al., 2024, *MNRAS*, 532, 2016
- McInnes L., Healy J., Melville J., 2018, arXiv e-prints
- McInnes L., Healy J., Saul N., Grossberger L., 2018, *The Journal of Open Source Software*, 3, 861
- McKee C. F., Ostriker J. P., 1977, *ApJ*, 218, 148
- McKee C. F., Storey J. W. V., Watson D. M., Green S., 1982, *ApJ*, 259, 647
- McKinney J., Cooper O. R., Casey C. M., Muñoz J. B., Akins H., Lambides E., Long A. S., 2025, *ApJ*, 985, L21
- Michel-Dansac L., Blaizot J., Garel T., Verhamme A., Kimm T., Trebitsch M., 2020, *A&A*, 635, A154
- Mingozi M. et al., 2022, *ApJ*, 939, 110
- Naidu R. P. et al., 2025, arXiv e-prints, arXiv:2505.11263
- Nakajima K. et al., 2025, arXiv e-prints, arXiv:2506.11846
- Nakazato Y., Yoshida N., Ceverino D., 2023, *ApJ*, 953, 140
- Nelson D. et al., 2018, *MNRAS*, 475, 624
- Nelson R. P., Langer W. D., 1997, *ApJ*, 482, 796
- Nomoto K., Kobayashi C., Tominaga N., 2013, *ARA&A*, 51, 457
- Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, *Nucl. Phys. A*, 777, 424
- Nyhaven C. T., Schimek A., Ciccone C., Decataldo D., Shen S., 2024, arXiv e-prints, arXiv:2410.18471
- Oh S. P., Haiman Z., Rees M. J., 2001, *ApJ*, 553, 73
- Omukai K., Tsuribe T., Schneider R., Ferrara A., 2005, *ApJ*, 626, 627
- Oppenheimer B. D., Schaye J., 2013, *MNRAS*, 434, 1043
- O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, *ApJ*, 807, L12
- Osterbrock D. E., Ferland G. J., 2006, *Astrophysics of gaseous nebulae and active galactic nuclei*
- Padoan P., Nordlund A., 2011, *ApJ*, 730, 40
- Pallottini A. et al., 2022, *MNRAS*, 513, 5621
- Pallottini A., Ferrara A., Gallerani S., Salvadori S., D'Odorico V., 2014, *MNRAS*, 440, 2498
- Park M., Conroy C., Johnson B. D., Leja J., Dotter A., Cargile P. A., 2024, arXiv e-prints, arXiv:2410.21375
- Pentericci L. et al., 2016, *ApJ*, 829, L11
- Pillepich A. et al., 2018, *MNRAS*, 475, 648
- Planck Collaboration et al., 2016, *A&A*, 594, A13
- Ploekinger S., Richings A. J., Schaye J., Trayford J. W., Schaller M., Chaikin E., 2025, arXiv e-prints, arXiv:2506.15773
- Pontzen A., Rey M. P., Cadiou C., Agertz O., Teyssier R., Read J., Orkney M. D. A., 2021, *MNRAS*, 501, 1755
- Pontzen A. et al., 2023, pynbody/pynbody: Version 1.5.2
- Pontzen A., Roškar R., Stinson G., Woods R., 2013, pynbody: N-Body/SPH analysis for python. Astrophysics Source Code Library, record ascl:1305.002
- Pontzen A., Tremmel M., Roth N., Peiris H. V., Saintonge A., Volonteri M., Quinn T., Governato F., 2017, *MNRAS*, 465, 547
- Ragan-Kelley M., Perez F., Granger B., Kluyver T., Ivanov P., Frederic J., Bussonnier M., 2014, in *AGU Fall Meeting Abstracts*, Vol. 2014, pp. H44D-07
- Raiter A., Schaerer D., Fosbury R. A. E., 2010, *A&A*, 523, A64
- Ramambason L. et al., 2022, *A&A*, 667, A35
- Reddy N. A. et al., 2018, *ApJ*, 869, 92
- Rémy-Ruyer A. et al., 2014, *A&A*, 563, A31
- Rey M., Katz H., Cadiou C., the MEGATRON team, 2025, arXiv e-prints
- Rey M. P. et al., 2023, *MNRAS*, 521, 995
- Rey M. P., Pontzen A., 2018, *MNRAS*, 474, 45
- Rey M. P., Pontzen A., Saintonge A., 2019, *MNRAS*, 485, 1906
- Rey M. P., Starkenburg T. K., 2022, *MNRAS*, 510, 4208
- Richings A. J., Faucher-Giguère C.-A., Gurvich A. B., Schaye J., Hayward C. C., 2022, *MNRAS*, 517, 1557
- Richings A. J., Schaye J., Oppenheimer B. D., 2014, *MNRAS*, 440, 3349
- Ritter C., Herwig F., Jones S., Pignatari M., Fryer C., Hirschi R., 2018, *MNRAS*, 480, 538
- Roberts-Borsani G. et al., 2024, *ApJ*, 976, 193
- Roberts-Borsani G. W., Ellis R. S., Laporte N., 2020, *MNRAS*, 497, 3440
- Röllig M., Osenkopf V., Jeyakumar S., Stutzki J., Sternberg A., 2006, *A&A*, 451, 917
- Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, *MNRAS*, 436, 2188
- Rosdahl J. et al., 2018, *MNRAS*, 479, 994
- Rosdahl J., Teyssier R., 2015, *MNRAS*, 449, 4380
- Roth N., Pontzen A., Peiris H. V., 2016, *MNRAS*, 455, 974
- Sanders R. L., Shapley A. E., Topping M. W., Reddy N. A., Brammer G. B., 2023, *ApJ*, 955, 54
- Sarmiento R., Scannapieco E., Cohen S., 2018, *ApJ*, 854, 75
- Saxena A. et al., 2024, arXiv e-prints, arXiv:2411.14532
- Saxena A. et al., 2021, *MNRAS*, 505, 4798
- Schaerer D., 2002, *A&A*, 382, 28
- Schaerer D., Guibert J., Marques-Chaves R., Martins F., 2025, *A&A*, 693, A271
- Schaerer D., Meynet G., Maeder A., Schaller G., 1993, *A&AS*, 98, 523
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schimek A., Ciccone C., Shen S., Decataldo D., Klaassen P., Mayer L., 2024, *A&A*, 687, L10
- Schmidt M., 1959, *ApJ*, 129, 243
- Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, *ApJ*, 571, 30
- Schreiber C. et al., 2015, *A&A*, 575, A74
- Seitenzahl I. R. et al., 2013, *MNRAS*, 429, 1156
- Senchyna P., Plat A., Stark D. P., Rudie G. C., Berg D., Charlot S., James B. L., Mingozi M., 2024, *ApJ*, 966, 92
- Shapley A. E. et al., 2019, *ApJ*, 881, L35
- Shapley A. E. et al., 2025, *ApJ*, 980, 242
- Shull J. M., van Steenberg M., 1982, *ApJS*, 48, 95
- Shuntov M. et al., 2025, *A&A*, 695, A20
- Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, 214, 15
- Stacy A., Bromm V., Lee A. T., 2016, *MNRAS*, 462, 1307
- Stancil P. C., Schultz D. R., Kimura M., Gu J. P., Hirsch G., Buenker R. J., 1999, *A&AS*, 140, 225
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Stanway E. R., Eldridge J. J., 2023, *MNRAS*, 522, 4430
- Stark D. P. et al., 2017, *MNRAS*, 464, 469
- Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, *ApJ*, 826, 159
- Sternberg A., Dalgarno A., 1989, *ApJ*, 338, 197
- Stopyra S., Pontzen A., Peiris H., Roth N., Rey M. P., 2021, *ApJS*, 252, 28

- Strait V. et al., 2023, ApJ, 949, L23
 Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, ApJ, 836, 164
 Suess K. A. et al., 2024, ApJ, 976, 101
 Teyssier R., 2002, A&A, 385, 337
 Tielens A. G. G. M., 2005, The Physics and Chemistry of the Interstellar Medium
 Topping M. W. et al., 2025, arXiv e-prints, arXiv:2502.08712
 Topping M. W. et al., 2024, MNRAS, 529, 4087
 Toro E. F., 2009, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction, 3rd edn. Springer, Berlin
 Trebitsch M. et al., 2021, A&A, 653, A154
 Trussler J. A. A. et al., 2025, MNRAS
 Trussler J. A. A. et al., 2023, MNRAS, 525, 5328
 Tumlinson J., Shull J. M., 2000, ApJ, 528, L65
 Tumlinson J., Shull J. M., Venkatesan A., 2003, ApJ, 584, 608
 Umeda H., Nomoto K., 2002, ApJ, 565, 385
 van der Walt S., Colbert S. C., Varoquaux G., 2011, Computing in Science and Engineering, 13, 22
 van Leer B., 1979, Journal of Computational Physics, 32, 101
 Vanzella E. et al., 2024, A&A, 691, A251
 Verner D. A., Ferland G. J., Korista K. T., Yakovlev D. G., 1996, ApJ, 465, 487
 Vikaeus A. et al., 2024, MNRAS, 529, 1299
 Virtanen P. et al., 2020, Nature Methods, 17, 261
 Vogelsberger M. et al., 2014, MNRAS, 444, 1518
 Vogelsberger M. et al., 2020, MNRAS, 492, 5167
 Voronov G. S., 1997, Atomic Data and Nuclear Data Tables, 65, 1
 Wang B. et al., 2024, ApJS, 270, 12
 Weaver J. R. et al., 2024, ApJS, 270, 7
 Weingartner J. C., Draine B. T., 2001, ApJ, 563, 842
 Wilkins S. M. et al., 2024, MNRAS, 527, 7965
 Wilkins S. M. et al., 2023, MNRAS, 522, 4014
 Willott C. J. et al., 2023, arXiv e-prints, arXiv:2311.12234
 Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, ApJ, 745, 50
 Witstok J. et al., 2024, arXiv e-prints, arXiv:2408.16608
 Wolfire M. G., Hollenbach D., McKee C. F., Tielens A. G. G. M., Bakes E. L. O., 1995, ApJ, 443, 152
 Wolfire M. G., McKee C. F., Hollenbach D., Tielens A. G. G. M., 2003, ApJ, 587, 278
 Xiao L., Stanway E. R., Eldridge J. J., 2018, MNRAS, 477, 904
 Xu H., Norman M. L., O’Shea B. W., Wise J. H., 2016, ApJ, 823, 140
 Yang H., Malhotra S., Rhoads J. E., Wang J., 2017, ApJ, 847, 38
 Yang S., Lidz A., Smith A., Benson A., Li H., 2023, MNRAS, 525, 5989
 Zel’dovich Y. B., 1970, A&A, 5, 84
 Zubko V., Dwek E., Arendt R. G., 2004, ApJS, 152, 211

APPENDIX

NUMERICAL METHODS

Here we provide a review of the primary numerical methods used for the MEGATRON simulations. Because the simulation methods are fully described in [Katz et al. \(2024c\)](#), here we only highlight the key galaxy formation physics, initial conditions, and additional developments.

Galaxy Formation Physics

Our simulations model gravity, radiation hydrodynamics, non-equilibrium thermochemistry, star formation, and feedback in a cosmological context. The gravitational potential is calculated on the AMR grid using a multigrid scheme ([Guillet & Teyssier 2011](#)). Hydrodynamics is evolved using a MUSCL-Hancock scheme ([van Leer 1979](#)), and the HLLC approximate Riemann solver ([Toro 2009](#)), assuming an adiabatic index of $\gamma = 5/3$ to close the relation between gas pressure and internal energy. Multi-frequency radiation transfer (RT) is solved in eight energy bins spanning the IR to the EUV using a two-moment (M1) scheme ([Rosdahl et al. 2013](#); [Levermore 1984](#)). We adopt a reduced speed of light approximation setting $c_{\text{sim}} = 0.01 c$. Radiation transport is subcycled up to 500 times per hydrodynamic time step ([Commerçon, Debout & Teyssier 2014](#)) and thermochemistry is computed during each subcycle. We model radiation pressure in the single scattering limit for UV and optical photons and multiple scatterings for IR photons ([Rosdahl & Teyssier 2015](#)).

Radiation and hydrodynamics are coupled to an 82 species non-equilibrium chemistry solver for primordial species, metals, and molecules. The solver follows H I-II, He I-III, e^- , C I-VI, N I-VII, O I-VIII, Ne I-X, Mg I-X, Si I-XI, S I-XI, Fe I-XI, H_2 , and CO. All metal ionization states that are not followed are assumed to be in collisional ionization equilibrium.

For each atomic species, we account for non-equilibrium ionization, recombination and charge exchange processes. Heating and cooling processes include photoheating, photoelectric heating, H_2 formation heating, H_2 excitation/dissociation heating, H_2 cooling, CO cooling, dust recombination cooling, dust-gas collisional processes, primordial heating/cooling, and metal line cooling. These processes are all modeled in non-equilibrium (see [Katz et al. 2022, 2024c](#) for the reaction network).

The reaction network and atomic data are derived from [Cen \(1992\)](#); [Hui & Gnedin \(1997\)](#); [Badnell \(2006\)](#); [Badnell et al. \(2003\)](#); [Aldrovandi & Pequignot \(1973\)](#); [Shull & van Steenberg \(1982\)](#); [Arnaud & Rothenflug \(1985\)](#); [Arnaud & Raymond \(1992\)](#); [Kingdon & Ferland \(1996\)](#); [Stancil et al. \(1999\)](#); [Barragán et al. \(2006\)](#); [Voronov \(1997\)](#); [Weingartner & Draine \(2001\)](#); [Verner et al. \(1996\)](#); [Glover et al. \(2010\)](#); [Gnedin, Tassis & Kravtsov \(2009\)](#); [Bialy & Sternberg \(2019\)](#); [Glover & Abel \(2008\)](#); [Baczynski, Glover & Klessen \(2015\)](#); [Nelson & Langer \(1997\)](#); [Glover & Clark \(2012\)](#); [Heays, Bosman & van Dishoeck \(2017\)](#). The dust-to-gas mass ratio as a function of metallicity is adopted following the empirical trends derived by [Rémy-Ruyer et al. \(2014\)](#) and we assume a dust composition following the BARE-GR-S model of [Zubko, Dwek & Arendt \(2004\)](#). Depletion of elements onto dust follows [Dopita et al. \(2000\)](#), and we adopt the solar abundance patterns of [Grevesse et al. \(2010\)](#) when converting to solar metallicity units. Heating and cooling rates are adopted from [Rosdahl et al. \(2013\)](#); [Bakes & Tielens \(1994\)](#); [Wolfire et al. \(1995, 2003\)](#); [Sternberg & Dalgarno \(1989\)](#); [Röllig et al. \(2006\)](#); [Black & Dalgarno \(1977\)](#); [Draine & Bertoldi \(1996\)](#); [Burton, Hollenbach & Tielens \(1990\)](#); [Cen \(1992\)](#); [Hui & Gnedin \(1997\)](#); [Osterbrock & Ferland \(2006\)](#); [Haiman, Thoul & Loeb \(1996\)](#); [Black \(1981\)](#); [McKee et al. \(1982\)](#); [Hollenbach & McKee \(1989\)](#); [Koyama & Inutsuka \(2000\)](#); [Draine \(2011\)](#); [Bialy & Sternberg \(2019\)](#); [Ferland et al. \(2017\)](#).

Star formation is modeled following a method that considers local turbulent properties of the gas ([Padoan & Nordlund 2011](#); [Federrath & Klessen 2012](#); [Kimm et al. 2017](#); [Rosdahl et al. 2018](#)). Gas cells can form stars if (1) their gas density

is $> 10 \text{ cm}^{-3}$ and $> 200\times$ the mean background, (2) the local turbulent Jeans length is unresolved¹⁵, (3) the gas cell represents a local maximum in the density field, and (4) the fluid flow is locally convergent. The star formation rate is then determined using a Schmidt law (Schmidt 1959), where the efficiency per free-fall time varies¹⁶ based on the local turbulent gas properties following results from high-resolution turbulent box simulations (Padoan & Nordlund 2011; Federrath & Klessen 2012).

The number of star particles formed is drawn from a Poisson distribution. When the metallicity of the gas is above the critical metallicity¹⁷ ($Z \geq 2 \times 10^{-8}$) Pop. II star formation occurs. We employ a minimum mass of $500 M_{\odot}$ for the high-redshift simulations¹⁸ or $4,600 M_{\odot}$ for the cosmic noon simulations¹⁹. Below the critical metallicity threshold, Pop. III star formation is modeled with individual stars drawn from a log-normal distribution with a mean mass of $100 M_{\odot}$ (Wise et al. 2012; Kimm et al. 2017). For all simulations, we assume a Kroupa IMF (Kroupa 2001) with a maximum mass of $300 M_{\odot}$. For the variable IMF simulation, we assume that the upper mass slope varies with gas density and metallicity following Marks et al. (2012).

Star particles interact with the gas via the radiation they emit, injecting energy and momentum in the form of feedback, and depositing mass and heavy elements into the gas. Stellar SEDs are adopted from the BPASS v2.2.1 binary stellar evolution model (Eldridge et al. 2017; Stanway & Eldridge 2018) for Pop. II stars and Schaerer (2002) for Pop. III stars²⁰. The Variable IMF simulation adopts SEDs that were computed with Starburst99 (Leitherer et al. 1999).

For Pop. II stars, we model core-collapse SN, type Ia SN, and stellar winds closely following Agertz et al. (2021); Agertz & Kravtsov (2015); Agertz et al. (2013), which was calibrated for Milky Way mass galaxies at $z = 0$, with only minor modifications. Progenitors of core-collapse SN are stochastically sampled from the IMF following the main-sequence lifetime-mass relation from Schaerer et al. (1993). Stars with $8 \leq m_{\star}/M_{\odot} \leq 25$ explode with an energy of $E_{\text{SN}} = 10^{51}$ erg. For the Variable IMF and HN, High ϵ_{ff} simulations, stars more massive than $25 M_{\odot}$ can explode as hypernovae Nomoto et al. (2006). In this case the energy depends on mass. The fraction of high-mass stars that explode as hypernovae strongly depends on metallicity (Kobayashi et al. 2006). SNIa explode following a delay time distribution of Maoz & Graur (2017), also injecting 10^{51} erg. In contrast to SN explosions, energy and momentum injection from stellar winds of lower-mass stars is computed as an IMF average (see Agertz et al. 2013).

Pop. III star particles either explode as core-collapse SN injecting $E_{\text{SN}} = 10^{51}$ erg for masses in the range $10 - 20 M_{\odot}$, explode as a hypernova in the mass range $20 - 40 M_{\odot}$, directly collapse to a black hole in the mass ranges $40 - 140 M_{\odot}$ and $> 300 M_{\odot}$, or explode as a PISN following Heger & Woosley (2002).

For both Pop. III and Pop. II SN, we check whether the cooling radius is resolved and then either inject energy or momentum. This method is calibrated so that the terminal momentum of SN explosions expected from high-resolution simulations (Kim & Ostriker 2015) is recovered independent of resolution and local gas conditions (Katz et al. 2024c).

Heavy elements are injected into the local oct of the star particle for SN (based on stochastic IMF sampling) and stellar winds (IMF averages) based on Limongi & Chieffi (2018); Ritter et al. (2018); Seitenzahl et al. (2013); Umeda & Nomoto (2002); Nomoto et al. (2006); Nomoto, Kobayashi & Tominaga (2013); Heger & Woosley (2002).

The appeal of such a detailed non-equilibrium thermochemistry and star formation is to naturally predict the (intrinsic) spectra of galaxies including the stellar continuum, nebular continuum, and nebular emission lines. For the stellar continuum, we simply adopt the same SED models that were used in the simulation, while emissivities for gas emission follow Ferland et al. (2017); Luridiana, Morisset & Shaw (2015); Dere et al. (2019). However, in certain cases, the H II regions around star particles are numerically unresolved. This typically happens for older star particles that have low ionizing output (and thus rarely contribute meaningfully to the spectrum), or very young stars that are embedded in dense gas clouds. When this occurs, the local cell properties become a mix between the neutral and ionized phases and must be corrected. We thus replace the emission of cells with unresolved Strömgren spheres by a spherical CLOUDY (Ferland et al. 2017) calculation adopting the gas properties of the host cell and SED of the star particle. We emphasize that in the spherical Strömgren sphere limit, CLOUDY and RAMSES-RTZ predict very similar emission line luminosities (to better than 10% on average).

We allow the AMR grid to refine on multiple criteria. A quasi-Lagrangian refinement strategy is employed so that a cell is refined into eight children cells when its dark matter or baryonic mass are equal to eight times that on the base grid in the initial conditions. Moreover, we ensure that the local Jeans length is resolved by at least four cells. For simulations that are part of our high-redshift suite, we allow for a constant-comoving spatial resolution such that the grid can refine up to the maximum level at any redshift. This corresponds to a resolution $1.7 \text{ pc } h^{-1}$ when the first Pop. III stars form at $z \sim 27$ and $5 \text{ pc } h^{-1}$ at $z = 8.5$. For the simulations primarily focused on studying the CGM towards Cosmic Noon, we employ constant physical refinement such that additional levels in the AMR grid are released at a fixed scale factor to maintain close to $24 \text{ pc } h^{-1}$ resolution at all times.

¹⁵ In most simulations, we compare the turbulent Jeans length to Δx but in the HN, High ϵ_{ff} simulation we consider $4\Delta x$ which encourages star formation at lower densities.

¹⁶ In the HN, High ϵ_{ff} simulation, the efficiency is always set to 100%.

¹⁷ There is some debate on the exact critical metallicity and it is unlikely to be a single number because physics such as redshift (via the CMB temperature), local radiation field, and dust probably all play a role (Bromm et al. 2001; Schneider et al. 2002; Omukai et al. 2005; Jappsen et al. 2007). We have adopted a very conservative value which assumes that dust cooling at low-metallicity can lead to fragmentation.

¹⁸ In the ‘HN, High ϵ_{ff} ’ simulation, the minimum stellar particle mass is $2,000 M_{\odot}$.

¹⁹ Note that the cosmic noon simulations begin with a metallicity floor and Pop. III star formation is not modeled.

²⁰ Note that Schaerer (2002) does not provide the full SED for Pop. III stars so when computing their spectra, we use the models of Larkin, Gerasimov & Burgasser (2023) for the same effective temperature.

Tracer Particles

In order to be able to track inflows and outflows of gas (and more generally any kinematics), we sample the Lagrangian history of the gas using the tracer particle implementation of [Cadiou, Dubois & Pichon \(2019\)](#).

Briefly, the passive tracers rely on a Monte-Carlo approach to sample mass fluxes between gas cells as well as from any gas cell onto stars (including Pop. II and Pop. III star formation) and from stars back into gas cells (including AGB winds and SNe). Mass fluxes between cells are sampled by moving particles across the cell boundaries with a probability $\Delta m_{\text{cell}}/m_{\text{cell}}$, where m_{cell} is the old cell mass and Δm_{cell} is the mass change of the cell as computed by the hydro solver. We also follow star formation in a similar fashion by attaching gas tracers onto newly-formed stars with probability $m_{\star}/m_{\text{cell}}$, where $m_{\star}$ is the mass of the star. Star tracers remain attached to their host star unless they are yielded back to the nearest gas cell in a wind or supernova event. This happens with probability $\Delta m_{\star}/m_{\star}$, where $\Delta m_{\star}$ is the mass lost by the star particle through either wind or supernova ejecta.

We initialize tracers in the zoomed region such that each initial cell contains $m_{\text{cell}}/m_{\text{tracer}}$ tracers (on average). Our tracer mass is $m_{\text{tracer}} = 7.3 \times 10^4 M_{\odot}$ ($150\times$ the stellar particle mass for most high-redshift simulations) for a total of 14,607,238 tracers. Tracer particles record the hydrodynamic properties of each gas cell. We also keep track of the number of times a tracer has been processed through a star and yielded back into the gas phase (either through winds or SN explosion).

Halo Finding

Halo es are extracted from the simulations using the ROCKSTAR 6D, phase-space, halo finder ([Behroozi, Wechsler & Wu 2013](#)). We only consider halo es with at least 300 high-resolution DM particles in the friends-of-friends group (before unbinding) and set the virial radius to be that where the enclosed average density is 200 times the mean background. The halo finder is run only on dark matter particles and when computing stellar and gas properties, we consider only those particles and cells that reside within 25% of the virial radius. This value was determined to maximize the total fraction of star particles assigned to halo es and minimize the overlap between main halo es and subhalo es. We empirically find there is $\leq 2\%$ overlap in terms of star particles assigned to more than one halo while nearly 95% of star particles have a host²¹. When computing spectra, we only consider halo es with at least one star particle, and those that are resolved by $>1,000$ dark matter particles and $>1,000$ gas cells, as this constitutes our “well-resolved” galaxy sample.

Dust Radiation Transport

Although in this work we focus primary on the intrinsic properties of galaxies (i.e. those prior to dust attenuation), absorption and scattering by dust can have significant impacts on the observable properties of high-redshift galaxies. For this reason, we have developed a custom interface²² to the public Monte Carlo radiation transport code RASCAS ([Michel-Dansac et al. 2020](#)). We adopt the same dust model that is used on-the-fly in the simulations that combines the empirical results for the dust-to-gas mass ratio as a function of metallicity ([Rémy-Ruyer et al. 2014](#)) with the BARE-GR-S dust composition from [Zubko, Dwek & Arendt \(2004\)](#). An example image combining multiple JWST filters into a single RGB image is shown in Figure 2. In addition to imaging, RASCAS also allows us to compute the dust-attenuated spectrum and a mock IFU along any sight line as also shown in Figure 2 for the massive $z = 4$ galaxy from the Cosmic Noon suite.

This paper was built using the Open Journal of Astrophysics L^AT_EX template. The OJA is a journal which provides fast and easy peer review for new papers in the `astro-ph` section of the arXiv, making the reviewing process simpler for authors and referees alike. Learn more at <http://astro.theoj.org>.

²¹ Part of the reason 100% of star particles are not assigned to halo es is because some star particles reside at larger distances than 25% of the virial radius while others form in halo es with fewer than 300 dark matter particles.

²² https://git-cral.univ-lyon1.fr/rascas/rascas/-/tree/megatron?ref_type=heads