







EDGE: the mass–metallicity relation as a critical test of galaxy formation physics

Oscar Agertz ¹★, Andrew Pontzen,² Justin I. Read ³, Martin P. Rey,²
Matthew Orkney,³ Joakim Rosdahl ⁴, Romain Teyssier ⁵, Robbert Verbeke,⁵
Michael Kretschmer ⁵ and Sarah Nickerson ⁵

¹Lund Observatory, Department of Astronomy and Theoretical Physics, Lund University, Box 43, SE-221 00 Lund, Sweden

²Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

³Department of Physics, University of Surrey, Guildford GU2 7XH, UK

⁴Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France

⁵Institute for Computational Science, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Accepted 2019 October 21. Received 2019 September 21; in original form 2019 April 4

ABSTRACT

We introduce the ‘Engineering Dwarfs at Galaxy Formation’s Edge’ (EDGE) project to study the cosmological formation and evolution of the smallest galaxies in the Universe. In this first paper, we explore the effects of resolution and sub-grid physics on a single low-mass halo ($M_{\text{halo}} = 10^9 M_{\odot}$), simulated to redshift $z = 0$ at a mass and spatial resolution of $\sim 20 M_{\odot}$ and ~ 3 pc. We consider different star formation prescriptions, supernova feedback strengths, and on-the-fly radiative transfer (RT). We show that RT changes the mode of galactic self-regulation at this halo mass, suppressing star formation by causing the interstellar and circumgalactic gas to remain predominantly warm ($\sim 10^4$ K) even before cosmic reionization. By contrast, without RT, star formation regulation occurs only through starbursts and their associated vigorous galactic outflows. In spite of this difference, the entire simulation suite (with the exception of models without any feedback) matches observed dwarf galaxy sizes, velocity dispersions, V -band magnitudes, and dynamical mass-to-light-ratios. This is because such structural scaling relations are predominantly set by the host dark matter halo, with the remaining model-to-model variation being smaller than the observational scatter. We find that only the stellar mass–metallicity relation differentiates the galaxy formation models. Explosive feedback ejects more metals from the dwarf, leading to a lower metallicity at a fixed stellar mass. We conclude that the stellar mass–metallicity relation of the very smallest galaxies provides a unique constraint on galaxy formation physics.

Key words: methods: numerical – galaxies: dwarf – galaxies: evolution – galaxies: formation – galaxies: kinematics and dynamics – Local Group.

1 INTRODUCTION

In our current Lambda cold dark matter (Λ CDM) cosmological paradigm (e.g. Komatsu et al. 2011; Planck Collaboration XVI 2014), galaxies form through the successive ‘hierarchical’ mergers of smaller galaxies (e.g. White & Rees 1978; Fall & Efstathiou 1980). This theory has been tremendously successful at matching the observed structure in the Universe on large scales (e.g. Springel, Frenk & White 2006; Vikhlinin et al. 2009). However, on smaller scales there have been long-standing tensions (e.g. Bullock & Boylan-Kolchin 2017). In part, such tensions owe to the uncertain

mapping between stars and dark matter (DM) in low-mass dwarf galaxies. This arises because dwarfs are very sensitive to the physics of galaxy formation. Isolated dwarfs can have their star formation suppressed or even extinguished by supernovae (SNe) feedback (e.g. Dekel & Silk 1986; Efstathiou 2000), reionization (e.g. Efstathiou 1992; Barkana & Loeb 1999; Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002), stellar winds and radiative feedback (e.g. Agertz et al. 2013; Hopkins et al. 2014), or some combination of these (e.g. Kravtsov, Gnedin & Klypin 2004; Read, Pontzen & Viel 2006b). If such dwarfs fall into a larger host galaxy, becoming satellites, then they are further buffeted by ram pressure stripping of their interstellar medium (ISM; e.g. Grebel, Gallagher & Harbeck 2003; Gatto et al. 2013) and tides (e.g. Kravtsov et al. 2004; Mayer et al. 2006; Read et al. 2006a).

* E-mail: oscar.agertz@astro.lu.se

The sensitivity of dwarfs to galaxy formation physics, like the gas density at which stars form (e.g. Kravtsov 2003; Saitoh et al. 2008) or the details of galactic outflows (e.g. Read et al. 2006b), makes them a natural ‘rosetta stone’ for constraining galaxy formation models. Early work simulating dwarf galaxies focused on high resolution small box simulations, stopping at high redshift ($z \sim 5-10$) to avoid gravitational collapse on the scale of the box (e.g. Abel et al. 1998; Gnedin & Kravtsov 2006; Read et al. 2006b; Mashchenko, Wadsley & Couchman 2008). These simulations demonstrated that stellar winds, SN feedback, and ionizing radiation combine to prevent star formation in halo masses below $\sim 10^{7-8} M_{\odot}$ (Read et al. 2006b; Bland-Hawthorn, Sutherland & Webster 2015). Furthermore, once cooling below 10^4 K is permitted and gas is allowed to reach high densities $n_{\max} \gtrsim 10 \text{ cm}^{-3}$ (requiring a spatial and mass resolution better than $\Delta x \lesssim 100 \text{ pc}$ and $m_{\min} < 10^3 M_{\odot}$; Pontzen & Governato 2012; Benitez-Llambay et al. 2019; Bose et al. 2019; Dutton et al. 2019), star formation becomes much more stochastic and violent (Mashchenko et al. 2008; Pontzen & Governato 2012; Dutton et al. 2019). The repeated action of gas cooling and blow-out due to feedback expels DM from the galaxy centre, transforming an initially dense DM cusp to a core (e.g. Read & Gilmore 2005; Pontzen & Governato 2012). Finally, independent of any internal sources of feedback energy, cosmic reionization can halt star formation in low-mass galaxies (e.g. Gnedin & Kravtsov 2006). Dwarfs that have not reached a mass of $M_{200} \gtrsim 10^8 M_{\odot}$ (see Section 2.4 for a definition of M_{200}) by the redshift that reionization begins ($z \sim 8-10$; Gnedin & Kaurov 2014; Ocvirk et al. 2018) are gradually starved of fresh cold gas, causing their star formation to shut down by a redshift of $z \sim 4$ (Oñorbe et al. 2015). This is similar to the age of nearby ‘ultra-faint’ dwarf galaxies (UFDs) that have $M_* \lesssim 10^5 M_{\odot}$, suggesting that at least some of these are likely to be relics from reionization, inhabiting pre-infall halo masses in the range $M_{200} \sim 10^{8-9} M_{\odot}$ (Gnedin & Kravtsov 2006; Bovill & Ricotti 2009, 2011; Brown et al. 2014; Weisz et al. 2014; Jethwa, Erkal & Belokurov 2018; Read & Erkal 2019).

To compare quantitatively to such very small galaxies in the local Universe (e.g. Simon 2019), we need simulations with sufficiently high resolution to model the ISM accurately. Most recent work in this area has focused on isolated dwarfs, removing the need to simultaneously capture a large host galaxy like the Milky Way. While much work has recently been done on simulating the smallest dwarf galaxies, the results from different groups, each of whom make different choices for their sub-grid physics model, box size, hydrodynamic and gravity solver, and resolution, are highly discrepant below $M_{200} \sim 10^{10} M_{\odot}$ (Oñorbe et al. 2015; Sawala et al. 2015; Wheeler et al. 2015; Read, Agertz & Collins 2016a; Fitts et al. 2017; Macciò et al. 2017; Munshi et al. 2017, 2019; Read et al. 2017; Revaz & Jablonka 2018; Wheeler et al. 2019; Smith, Sijacki & Shen 2019). While many simulations reproduce the observed structural properties of dwarf galaxies, such as galaxy sizes and velocity dispersions, there are orders of magnitude differences in the stellar mass to halo mass relation, and simulations struggle to reproduce the stellar mass–metallicity relation of the faintest dwarfs (see discussion below; and for a data compilation, see Figs 3 and 6). Furthermore, none of the simulations has yet managed to produce a fully convincing explanation for the existence of *star-forming* ultra-faint dwarf galaxies like Leo T (except possibly Wright et al. 2019, see also Verbeke, Vandenbroucke & Rijke 2015), as cosmic reionization ought to have evaporated all cold gas from such low-mass galaxies. Finally, while observationally there is mounting evidence for DM cores in at least some UFDs (Amorisco 2017; Contenta et al. 2018; Sanders, Evans & Dehnen 2018), some groups

find core formation can occur ‘all the way down’ to the very lowest mass dwarfs (Read et al. 2016a; Munshi et al. 2017), while others find that core formation ceases below $\sim 10^{10} M_{\odot}$ (Oñorbe et al. 2015; Tollet et al. 2016).

In this paper, we introduce a new dwarf galaxy simulation campaign: ‘Engineering Dwarfs at Galaxy Formation’s Edge’ (EDGE) with the goal of shedding light on the above discrepancies. There are several new elements to EDGE that make this investigation possible. First, we work at a mass and spatial resolution of $\sim 20 M_{\odot}$ and $\sim 3 \text{ pc}$, respectively, allowing us to better capture the impact of each individual SN explosion on the forming dwarf. This simplifies the sub-grid modelling of the SN, reducing the need for delayed cooling, momentum capturing schemes, or other similar prescriptions that are required at lower resolution to prevent overcooling of the SN-heated gas (see e.g. Stinson et al. 2006; Dalla Vecchia & Schaye 2008; Torrey et al. 2014; Crain et al. 2015 and for a discussion, see Read et al. 2016a). Secondly, we explore the effect of switching on each piece of our sub-grid model for star formation and feedback one at a time, allowing us to assess its role in shaping the final observed properties of the dwarf today. Thirdly, we model isolated dwarfs in a void region down to redshift zero, using a zoom technique. This allows us to explicitly compare our results with observations of ultra-faint dwarfs that can only be seen, at present, in the Local Group. Finally, we set up our cosmological initial conditions using the new GENETIC code (Roth, Pontzen & Peiris 2016; Rey & Pontzen 2018). This will allow us (in forthcoming papers) to forensically explore the effect of different merger histories and environments on the properties of dwarf galaxies. In this first paper in the series, we present the results of 16 hydrodynamical simulations of a single $M_{200} = 10^9 M_{\odot}$ dwarf run at different resolutions and with different sub-grid physics models. In particular, we explore different star formation prescriptions, the effect of SN feedback, and the effect of on-the-fly radiative transfer (RT). We emphasize that our goal in running this large suite of simulations is not to ‘calibrate’ our sub-grid physics model. Rather, we seek to address the following questions: (i) which observables are most sensitive to changes in our sub-grid model? and (ii) which physics are most important for regulating star formation and determining the final observed properties of the smallest dwarfs? In further papers in the series, we will explore the role of physics not considered in this work, the role of different merger histories, the effect of increasing halo mass, and the impact of star formation and feedback on the inner density of the dwarf’s DM halo.

This paper is organized as follows. In Section 2, we describe the code that we use to run the simulations, the sub-grid physics models that we explore in this work, and how we set up the initial conditions. In Section 3, we present mass growth histories, the stellar mass–halo mass relation, and scaling relations for all simulated dwarf galaxies. In Section 4, we compare our findings to observations and previous simulations in the literature and outline the limitations of our models. Finally, in Section 5, we present our conclusions.

2 SIMULATIONS

We use RAMSES-RT (Rosdahl et al. 2013; Rosdahl & Teyssier 2015), which is a radiation hydrodynamics (RHD) extension of the cosmological Adaptive Mesh Refinement (AMR) hydrodynamical code RAMSES (Teyssier 2002),¹ to solve the evolution of DM, stellar

¹The public code, including all the RHD extensions used here, can be downloaded at <https://bitbucket.org/rteyssie/ramses>.

populations, and gas via gravity, hydrodynamics, RT, and non-equilibrium radiative cooling/heating. For hydrodynamics, we use the HLLC Riemann solver (Toro, Spruce & Speares 1994) and the MinMod slope limiter to construct gas variables at cell interfaces from their cell-centred values. To close the relation between gas pressure and internal energy, we use an ideal gas equation of state with an adiabatic index $\gamma = 5/3$. The dynamics of collisionless DM and star particles are evolved with a multigrid particle-mesh solver and cloud-in-cell interpolation (Guillet & Teyssier 2011). The advection of radiation between cells is solved with a first-order moment method, using the fully local M1 closure for the Eddington tensor (Levermore 1984) and the Global-Lax-Friedrich flux function for constructing the inter-cell radiation field. With the M1 closure, the collisionless nature of photons is lost and beams are not perfectly maintained (see e.g. Rosdahl et al. 2013). This has the effect that in the case where radiation from many sources is mixed, the radiation flux becomes distorted, by up to a factor two (Decataldo et al. 2019), compared to the real flux which can for example be obtained with ray-tracing methods. These more accurate methods, however, are prohibitively expensive to use in simulations with the number of resolution elements and, more importantly, number of radiation sources that we model in this work. We argue that the modest factor of two errors produced by the M1 closure are likely less significant than several other approximations going into our galaxy formation models.

In the following subsections, we describe the set-up of our simulations (multifrequency radiation, thermochemistry, initial conditions) and the adopted galaxy formation physics (star formation and stellar feedback).

2.1 Radiation

Star particles (see 2.3) are treated as single stellar populations (SSPs) with spectral energy distributions (SEDs) taken from Bruzual & Charlot (2003) (see Rosdahl et al. 2013, for details). We employ six photon groups to account for (1) photoionization heating, where three groups bracket the ionization energies for H I, He I, and He II, (2) H₂ dissociating Lyman–Werner radiation, (3) direct (single scattering) radiation pressure,² and (4) non-thermal radiation pressure from multiscattered IR photons. For dust opacities, we adopt $\kappa = 10 (Z/Z_{\odot}) \text{ cm}^2 \text{ g}^{-1}$ for the IR photons, while for the higher energy photons we assume $\kappa = 1000 (Z/Z_{\odot}) \text{ cm}^2 \text{ g}^{-1}$, with Z/Z_{\odot} being the gas metallicity in units of the solar value (here taken to be $Z_{\odot} = 0.02$). For a full description of our treatment of dust physics, see Rosdahl & Teyssier (2015) and Kimm et al. (2017). Although we include a treatment of radiation pressure in our RT scheme we find that it has negligible impact, likely due to the low dust content in the simulated UFDs; final galaxy stellar masses in two tests without any radiation pressure (not presented here for brevity) end up within ± 20 per cent of stellar masses in models including it (but see Wise et al. 2012a).

The group properties (average energies and cross-sections to molecular hydrogen, hydrogen, and helium) are updated every 10 coarse time-steps from luminosity-weighted averages of the spectra of all stellar populations in the simulation volume, as described in Rosdahl et al. (2013). We do this so that at any time, the cross-sections and photon energies are representative of the luminous stellar populations.

²adopting the ‘reduced flux approximation’ described in Appendix B of Rosdahl et al. (2015).

In Table 1 we summarize the above information, and present the average group energies and cross-sections, over the entire 13.7 Gyr simulation time from one of our high-resolution simulations discussed below.³ The values only change by a few tens of per cent over the course of a simulation as the contributions from both old and metal-rich stellar populations increase.

2.2 Gas thermochemistry

For $T > 10^4$ K, the contribution from metals to gas cooling is computed using tables generated with CLOUDY (Ferland et al. 1998, version 6.02). We adopt the UV background of Faucher-Giguère et al. (2009) as our standard setting, but also consider that of Haardt & Madau (1996) (see Section 2.5). The homogenous UV background is necessary to include as we are not capturing radiation sources outside the zoom region. It is not treated using RT but in the commonly adopted optically thin, cell-by-cell heating approximation (e.g. Katz, Weinberg & Hernquist 1996). Following Rosdahl & Blaizot (2012) self-shielding against the homogenous UV background is modelled by applying a damping function to the photoionization rate, $\Gamma_{\text{ss}} = \Gamma_{\text{UV}} \exp(-n_{\text{H}}/10^{-2} \text{ cm}^{-3})$, for hydrogen densities $n_{\text{H}} > 10^{-2} \text{ cm}^{-3}$. Self-shielding against radiation modelled using RT is treated self-consistently.

For $T \leq 10^4$ K, we use the fine structure cooling rates from Rosen & Bregman (1995). The non-equilibrium hydrogen and helium thermochemistry, coupled with the local ionizing radiation, is performed with the quasi-implicit method described in Rosdahl et al. (2013) via photoionization, collisional ionization, collisional excitation, recombination, bremsstrahlung, homogeneous Compton cooling/heating off the cosmic-microwave background, and dielectronic recombination. We account for the formation, advection, destruction, and cooling of molecular hydrogen (H₂) (see Nickerson, Teyssier & Rosdahl 2018, for details), and its coupling to the radiation field, in all simulations unless otherwise stated.

Along with the temperature, and photon fluxes, we track, in every cell, the fractions of neutral hydrogen, ionized hydrogen, singly, and doubly ionized helium ($x_{\text{H I}}$, $x_{\text{H II}}$, $x_{\text{He II}}$, $x_{\text{He III}}$, respectively), and advect them with the gas as passive scalars. The thermochemistry is operator split from the advection of gas and radiation and performed with adaptive time-step sub-cycling on every RT time-step. To keep the computational costs low, we use a reduced speed of light, $\bar{c} = 10^{-2} c$ (Rosdahl et al. 2013), where c is the true speed of light.

2.3 Galaxy formation physics

Star formation follows a Schmidt law,

$$\dot{\rho}_{*} = \epsilon_{\text{ff}} \frac{\rho_{\text{g}}}{t_{\text{ff}}} \text{ for } \rho_{\text{g}} > \rho_{*}, \quad (1)$$

where ρ_{g} is the gas density, ρ_{*} the density threshold of star formation, $t_{\text{ff}} = \sqrt{3\pi/32G\rho}$ is the local gas free-fall time, and ϵ_{ff} is the star formation efficiency per free-fall time. We adopt $\rho_{*} = 300 m_{\text{H}} \text{ cm}^{-3}$, and sample equation (1) stochastically on a cell-by-cell basis at every fine simulation time-step using 300 M_{\odot} star particles⁴ (see Agertz et al. 2013). Furthermore, we only allow stars to form from cold gas ($T < 100$ K). In a subset of simulations,

³Similar values are obtained from all of our simulations where RT is included.

⁴Chosen to accommodate the (discrete and stochastic) mass-loss from several SN type II explosions, which is not the case if initial star particle masses are set too close to the mass of individual massive stars in our current

Table 1. Photon group energy (frequency) intervals and properties from the ‘Hires + RT’ simulation. The energy intervals defined by the groups are indicated in units of eV by ϵ_0 and ϵ_1 . The last four columns show photon properties derived every 10 coarse time-steps from the stellar-luminosity-weighted SED model. These properties evolve over time as the stellar populations age, and the mean values are quoted. $\bar{\epsilon}$ denotes the photon energies, while σ_{H_2} , $\sigma_{\text{H I}}$, $\sigma_{\text{He I}}$, and $\sigma_{\text{He II}}$ denote the cross-sections for ionization of molecular hydrogen, hydrogen, and helium, respectively.

Photon group	ϵ_0 [eV]	ϵ_1 [eV]	$\bar{\epsilon}$ [eV]	σ_{H_2} [cm ²]	$\sigma_{\text{H I}}$ [cm ²]	$\sigma_{\text{He I}}$ [cm ²]	$\sigma_{\text{He II}}$ [cm ²]
IR	0.1	1.0	0.6	0	0	0	0
Optical + FUV	1.0	12.0	3.2	0	0	0	0
Lyman–Werner	12.0	13.6	12.6	1.7×10^{-19}	0	0	0
UV _{H I}	13.6	24.59	18.1	5.1×10^{-18}	3.3×10^{-18}	0	0
UV _{He I}	24.59	54.42	35.6	2.0×10^{-18}	5.9×10^{-19}	4.3×10^{-18}	0
UV _{He II}	54.42	∞	64.9	3.4×10^{-19}	8.3×10^{-20}	1.2×10^{-18}	1.1×10^{-18}

we require star formation to only occur in molecular gas, as motivated by the observed close to linear relation between Σ_{mol} and Σ_{SFR} (e.g. Bigiel et al. 2008). In those simulations, $\rho_g \rightarrow f_{\text{H}_2} \rho_g$ in equation (1), where f_{H_2} is the molecular hydrogen fraction in a cell (see also Gnedin, Tassis & Kravtsov 2009; Gnedin & Kravtsov 2010; Christensen et al. 2014). We note that while a correlation between molecular hydrogen and star formation is well motivated both theoretically and empirically, it is not well established in the low-metallicity regimes probed in this work⁵ ([Fe/H] $\lesssim -2$, Glover & Clark 2012; Krumholz 2012). For this reason we consider the use of H₂-based star formation models in the UFD regime as exploratory (see also Munshi et al. 2019).

Observationally, ϵ_{ff} averages 1 per cent on galactic kpc scales (Bigiel et al. 2008) as well as in Milky Way giant molecular clouds (GMCs) (Krumholz & Tan 2007), albeit with a spread of several dex (Murray 2011; Lee, Miville-Deschênes & Murray 2016). Recently Grisdale et al. (2019) demonstrated how high efficiencies ($\epsilon_{\text{ff}} \sim 10$ per cent) on scales of parsecs, coupled to a feedback budget like the one adopted here, provide a close match to the observed (i.e. emerging) efficiencies on scales of individual GMCs. Motivated by these findings we adopt $\epsilon_{\text{ff}} = 10$ per cent.

We adopt the stellar feedback budget described in Agertz et al. (2013). Briefly, this feedback prescription includes the injection of energy, momentum, mass, and heavy elements over time from SNII and SNIa explosions and stellar winds into the surrounding ISM. In contrast to Agertz & Kravtsov (2015), we do not adopt a sub-grid model for radiation pressure, as this is self-consistently treated by the RT solver (see Section 2.1, and Section 2.5 for which simulations adopt RT). Each mechanism depends on the stellar age, mass, and gas/stellar metallicity, calibrated on the stellar evolution code STARBURST99 (Leitherer et al. 1999), treating each formed stellar particle as a single-age stellar population with a Kroupa (2001) initial mass function (IMF). Feedback is done continuously at the appropriate times when each feedback process is known to operate, taking into account the lifetime of stars of different masses within a stellar population.

We track iron (Fe) and oxygen (O) abundances separately, and advect them as passive scalars. When computing the gas cooling

feedback scheme. Note also that star particle masses, on average, are reduced by up to 50 per cent due to stellar evolution (e.g. Leitner 2012).

⁵The cooling and star formation time-scales can in metal-poor environments be shorter than the time-scale for reaching an equilibrium chemical state for which the gas would be H₂ dominated. Star formation can under such conditions correlate with atomic gas (e.g. Krumholz 2012). The non-equilibrium treatment of H₂ in our adopted method (Nickerson et al. 2018) mitigates this issue (see also Krumholz & Gnedin 2011).

rate, which is a function of total metallicity, we construct a total metal mass as

$$M_Z = 2.09M_{\text{O}} + 1.06M_{\text{Fe}} \quad (2)$$

according to the solar abundances of alpha (C, N, O, Ne, Mg, Si, S) and iron (Fe, Ni) group elements of Asplund et al. (2009).

SNe explosions are modelled as discrete events, and we follow the approach by Kim & Ostriker (2015) (see also Martizzi, Faucher-Giguère & Quataert 2015) and inject the full momentum generated during the Sedov–Taylor phase if an SN cooling radius is not captured with at least six grid cells, otherwise we inject an $E_{\text{SN}} = 10^{51}$ erg of thermal energy (see Agertz, Romeo & Grisdale 2015, for details) and allow for the hydrodynamic solver to track the build-up of momentum. At the numerical resolution adopted here, > 90 per cent of all SN explosions are resolved by at least SIX cooling radii in our fiducial simulation. We note that we do not enforce any additional refinement criterion to achieve this; the (Lagrangian) mass-based refinement scheme, discussed in Section 2.5, is enough for this to be satisfied. This feedback budget has been shown to lead to Milky Way disc galaxies in close agreement to observations (Agerz & Kravtsov 2016, Agertz et al., in preparation), bursty star formation and realistic properties of dwarf galaxies (Read et al. 2016a, b), and an ISM and structure of GMCs in excellent agreement with observations (Grisdale et al. 2017, 2018).

Finally, in order to account for enrichment from unresolved Population III (Pop III) star formation, we adopt a pre-existing metal floor at $Z = 10^{-3}Z_{\odot}$ (e.g. Wise et al. 2012b; Jaacks et al. 2018) added to the oxygen field.⁶ We note however that H₂ cooling is the main coolant at $Z \lesssim 10^{-2}Z_{\odot}$, which is the regime we are exploring in this work. As soon as star formation and enrichment begins, metal line cooling, and subsequently atomic line cooling, also become important coolants (Wise et al. 2014). The inclusion of a Pop III floor therefore has a small effect on star formation properties in our simulations, and tests with $Z \leq 10^{-4}Z_{\odot}$ show that our choices have no impact on any of the conclusions presented in this paper.

2.4 Initial conditions

In this work, we study the cosmological formation of dwarf galaxies forming in isolation, hence removing complexities such as gas stripping during infall and environmental star formation quenching.

⁶Motivated by the abundance ratios for alpha group elements observed in extremely metal-poor stars (Iwamoto et al. 2005), but we note that Pop III yields are *extremely* uncertain, and that abundance ratios and the overall metal content is likely dependent on environment (e.g. Jaacks et al. 2018).

Being isolated, deformation due to tides is also minimized, and we can study the ‘pristine’ galaxy formation scenario in a $10^9 M_\odot$ DM halo. To generate initial conditions, we use the code GENETIC (Roth et al. 2016; Rey & Pontzen 2018) which will in future work allow us to explore a continuum of alternative merger histories for the same galaxy. For this paper, the most important capability of GENETIC is simply to recursively refine regions of the simulation on extremely fine grids. We first generated initial conditions for a simulation with box size of $L_{\text{box}} = 50$ Mpc with cosmological parameters $\Omega_m = 0.309$, $\Omega_\Lambda = 0.691$, $\Omega_b = 0.045$, and $H_0 = 67.77 \text{ km s}^{-1} \text{ Mpc}^{-1}$, in line with data from the PLANCK satellite (Planck Collaboration XVI 2014).

We simulated this volume with only DM from $z = 99$ to $z = 0$, using 512^3 resolution elements (giving a particle mass of $m_{\text{DM}} = 3.8 \times 10^7 M_\odot$). Next, we picked the largest void and resimulated it (again with only DM) at the equivalent of 2048^3 resolution ($m_{\text{DM}} = 4.9 \times 10^5 M_\odot$), adding the appropriate small scale power to this grid. We then identified DM haloes within the void at $z = 0$ using the HOP halo finder (Eisenstein & Hut 1998) and computed their virial masses, M_{200} , defined as the mass inside of a spherical volume encompassing 200 times the cosmic critical density $\rho_{\text{crit}} = 3H(z)^2/8\pi G$, where $H(z)$ is the Hubble parameter. The radial extent of this spherical volume defines the virial radius, r_{200} .

To find haloes in isolation, we computed pairwise distances, measured from the edges of their individual virial radii. For halo n and m , the distance is $D_{n,m} = |\mathbf{r}_n - \mathbf{r}_m| - r_{200,n} - r_{200,m}$, where \mathbf{r} is the position of a halo centre. Expressed in units of the virial radius of halo n , $I_{n,m} = D_{n,m}/r_{200,n}$. As a measure of isolation for halo n , we define an *isolation parameter* as the minimum of $I_{n,m}$, i.e. $I_n = \min(I_{n,m})$. At $z = 0$, we only considered haloes in the mass range $0.8 \times 10^9 M_\odot < M_{200} < 1.2 \times 10^9 M_\odot$, with $I_n > 10$ for halo pairs with $M_{200,m}/M_{200,n} > 1$.

From the filtered halo catalogue, we picked a halo that at $z = 0$ had a mass of $M_{200} \sim 10^9 M_\odot$ and was visually isolated from massive DM filaments. Having found a DM halo satisfying our selection criteria, we identified all particles belonging to this halo out to $2 \times r_{200}$ at $z = 0$ and traced the particles back to the initial conditions ($z = 99$). For the Lagrangian region of this halo, we generated separate new initial conditions at the equivalent of 16384^3 resolution ($m_{\text{DM}} = 1106 M_\odot$), again adding small-scale density fluctuations compatible with the background. These ‘fiducial’ initial conditions are then modified to include baryons as well as DM, forming the basis for the simulation suite described below. For the ‘Hires’ simulations, we further refined the Lagrangian region by a factor of eight in mass.

2.5 Simulation suite

For haloes of a given mass, there will be a diversity of cosmologically determined mass accretion histories and environments giving rise to a spread in final observed properties at $z = 0$. Our EDGE project will ultimately explore this diversity using ‘genetic modification’ (Roth et al. 2016; Rey & Pontzen 2018). However, before assessing the connection between history and observables, we need a firm handle on theoretical uncertainties resulting from the small-scale star formation and feedback physics. While there will never be a ‘complete’ account of feedback physics, understanding the leading-order uncertainties and their implications for interpreting future simulations is a key first step.

This first work therefore probes the effect of differing physics implementations, and we undertake this study at two numerical

resolutions. In the high resolution simulations, the DM particle resolution is $m_{\text{DM}} = 118 M_\odot$, and the equivalent baryon resolution (the mass of baryons at the finest grid level in the initial conditions) is $m_{\text{bar}} = \Omega_b/\Omega_m m_{\text{DM}} = 20 M_\odot$. For our fiducial resolution, the corresponding numbers are $m_{\text{DM}} = 945 M_\odot$ and $m_{\text{bar}} = 161 M_\odot$. In all simulations, we reach a mean physical resolution of $\Delta x = 3$ pc at all times in the inner parts of DM haloes and in the ISM. Refinement is based on a pseudo-Lagrangian approach, where a cell is split if its mass m_{cell} exceeds $8 \times m_{\text{bar}}$, where the cell mass accounts for both stars and gas. In addition, a cell is allowed to refine if it contains more than eight DM particles. All simulations are run from $a = 0.01$ ($z = 99$) to $a = 1$ ($z = 0$), and simulation snapshots are stored every $\Delta a = 0.01$.

For both resolution settings, we adopt the standard galaxy formation physics presented above and refer to these as ‘Fiducial + RT’ and ‘Hires + RT’, and ‘Fiducial’ and ‘Hires’ when RT is not included. We re-emphasize that without on-the-fly RT, we do not consider any *sub-grid* model of radiative feedback (e.g. Agertz et al. 2013), only feedback from stellar winds and SNe. For the fiducial resolution setting, we carry out a suite of simulations where the sensitivity to galaxy formation physics is tested as follows: increasing the strength of SN feedback ($E_{\text{SN}} = 2, 10$, and 100×10^{51}), not considering H_2 physics and the associated gas cooling, changing the UV background field (changing from the fiducial Faucher-Giguère et al. 2009 UV background to that of Haardt & Madau 1996), adopting an H_2 -based star formation model (see Section 2.3), and studying the impact of coupling it to RT. In the latter model, molecular hydrogen can be dissociated by Lyman–Werner radiation, forcing stars to only form in environments where H_2 is self-shielded, as well as shielded by dust.

For both resolution settings, with and without RT, we also carry out simulations where we artificially limit the efficiency of SN wind driving by limiting the maximum allowed temperatures of the gas to $T_{\text{max}} = 10^8$ K and restrict the maximum allowed velocities of feedback, upon injection, to $v_{\text{fb,max}} = 1000 \text{ km s}^{-1}$ (approximately leading to post-shock temperatures of $\sim 10^8$ K).⁷ We refer to these models as ‘weak feedback’. Although the exact values of such feedback limiters are arbitrary, these tests illustrate the role of hot fast-moving winds in regulating ultra-faint dwarf formation, as we will demonstrate below. We summarize the entire simulations suite, as well as the $z = 0$ properties of the central dwarf galaxy, in Table 2.

3 RESULTS

Fig. 1 shows the gas surface density, temperature, and metallicity in a $5 \times 5 \text{ kpc}^2$ region centred on the main dwarf galaxy in the ‘Hires’ and ‘Hires + RT’ simulations at $z = 10, 8, 6.5$, and 5. The maps illustrate the dramatic effect of stellar feedback, and how the inclusion of radiative feedback changes the mode of galaxy formation. In ‘Hires’, star formation leads to large-scale hot ($T > 10^6$ K) SN-driven outflows at all times before $z \sim 5$. The enriched winds efficiently mix into the intergalactic medium (IGM), leading to an average background $[\text{Fe}/\text{H}] > -3$ even at early times ($z \sim 10$).

This is in stark contrast to the RT counterpart, where the gas is kept warm ($T \sim 10^4$ K) and large scale metal rich outflows are relatively weak until $z \lesssim 6.5$. Before this epoch, the IGM is metal poor ($[\text{Fe}/\text{H}] \lesssim -4$), and metal line cooling is subdominant to H_2 cooling. At $z = 5$, reionization has been operating for ~ 500 Myr,

⁷In the adopted feedback model, SN ejecta from $8 M_\odot$ stars travel with velocities of $\sim 3000 \text{ km s}^{-1}$

Table 2. Simulations and their $z = 0$ properties. All simulations use the same cosmological zoom initial conditions, are run with RAMSES-RT and reach a minimum cell size of $\Delta x_{\min} = 3$ pc. The specified quantities are, from left to right: the DM particle mass in the deepest refinement region; the gas refinement mass above which new cells will be opened; the total stellar mass formed by $z = 0$; the V -band magnitude computed using SUNSET ray-tracing; the half-mass radius; the 1D equivalent velocity dispersion; the dynamical mass-to-light ratio within this radius; the enrichment relative to solar; and the mean star formation rate over the time until the galaxy quenches.

Simulation	m_{DM} [M_{\odot}]	m_{bar} [M_{\odot}]	M_{\star} [$10^5 M_{\odot}$]	M_V	$r_{1/2}$ [pc]	σ_{\star} [km s^{-1}]	M_{dyn}/L ($< r_{1/2}$)	[Fe/H] [dex]	$\langle \text{SFR} \rangle_a$ [$M_{\odot} \text{ yr}^{-1}$]
Fiducial, no feedback	945	161	172.1	−12.0	51	17.6	2	no enrichment	1.4×10^{-2}
Fiducial	945	161	1.2	−6.6	313	6.1	287	-2.65 ± 0.82	1.1×10^{-4}
Fiducial (weak feedback) ^b	945	161	15.8	−9.4	336	6.1	17	-1.07 ± 0.68	1.4×10^{-3}
Fiducial, $2 \times E_{\text{SN}}$	945	161	1.2	−6.6	280	6.3	212	-2.66 ± 0.79	1.1×10^{-4}
Fiducial, $10 \times E_{\text{SN}}$	945	161	0.52	−5.7	314	6.9	647	-3.42 ± 0.96	4.8×10^{-5}
Fiducial, $100 \times E_{\text{SN}}$	945	161	0.11	−4.0	157	5.6	949	-4.00 ± 1.20	2×10^{-4}
Fiducial, H ₂ SF	945	161	1.2	−6.6	285	6.1	189	-2.49 ± 0.99	1.1×10^{-4}
Fiducial, no H ₂ physics	945	161	0.53	−5.7	265	5.6	338	-2.27 ± 1.29	4.8×10^{-5}
Fiducial, HM UV	945	161	0.94	−6.35	283	6.6	287	-2.59 ± 0.79	9.4×10^{-5}
Fiducial+RT	945	161	0.30	−5.7	308	5.6	398	-2.31 ± 0.88	2.7×10^{-5}
Fiducial + RT (weak feedback) ^b	945	161	1.1	−7.1	412	6.5	200.8	-2.35 ± 0.96	9.7×10^{-5}
Fiducial + RT, H ₂ -based SF	945	161	0.62	−6.5	597	6.5	535	-1.91 ± 0.65	5.3×10^{-5}
Hires	118	20	2.5	−7.4	370	7.2	171	-2.61 ± 0.95	2.3×10^{-4}
Hires (weak feedback) ^b	118	20	20.5	−9.7	311	6.1	12.3	-1.01 ± 0.91	1.8×10^{-3}
Hires+RT	118	20	0.31	−5.7	203	5.2	226	-2.6 ± 1.37	2.9×10^{-5}
Hires + RT (weak feedback) ^b	118	20	2.8	−8.1	368	6.3	67.1	-1.32 ± 0.85	2.3×10^{-4}
Fiducial, DM only	1106	–	–	–	–	–	–	–	–
Hires, DM only	138	–	–	–	–	–	–	–	–

^aDefined as M_{\star}/t_{SF} , where t_{SF} is the duration of star formation for the galaxy, here ~ 1 –1.2 Gyr in all simulations.

^bMaximum allowed gas temperatures $T_{\text{max}} = 10^8$ K, maximum allowed SN and stellar winds velocities $v_{\text{fb,max}} = 10^3 \text{ km s}^{-1}$.

and large-scale filamentary structures have been evaporated, with stars still forming from the residual cold gas from earlier accretion epochs. At this time, both simulations feature a circumgalactic medium enriched to $[\text{Fe}/\text{H}] \sim -2$. ‘Hires’, due to its early intense outflows, retains an enriched IGM to larger radii than ‘Hires + RT’, and the RT simulation also features a visibly denser circumgalactic medium compared to ‘Hires’. In the next sections, we identify differences in the ISM that give rise to the modified outflow behaviours.

3.1 Mass growth histories

Fig. 2 shows the build-up of stellar and DM halo masses (M_{200}) for the simulations in Table 2. Stellar masses (M_{\star}) are throughout this paper defined as the total mass of stars in the inner part of the DM halo ($r < 0.25R_{200}$). The left-hand panel includes the models targeting RT and H₂ physics, while the right-hand panel focuses on variations in feedback strength and numerical resolution. Star formation starts at $z \sim 12$ –13 in all models, with delays introduced by variations in how H₂ physics is treated; see below. The virial temperature of the halo at $z < 10$ is $T_{\text{vir}} \sim 5000$ –9000 K, i.e. below the ‘atomic-cooling’ regime ($T_{\text{vir}} \gtrsim 10^4$ K). As reionization heats the gas to $T \sim 10^4$ K, this leads to a cessation of gas accretion and ultimately quenching of star formation. Star formation from the already cold and dense (self-shielded) gas in the ISM can, however, proceed throughout the epoch of reionization. Indeed, the last stars form around $z \sim 4$ –5 in all simulations.

Without any stellar feedback, the final stellar mass is close to $M_{\star} = 2 \times 10^7 M_{\odot}$. By including stellar feedback (but neglecting RT), stellar masses are lowered by over two orders of magnitude to $M_{\star} = 1.2 \times 10^5 M_{\odot}$ in ‘Fiducial’, and $\sim 2.5 \times 10^5 M_{\odot}$ in ‘Hires’. This result is independent of the adopted UV background, with ‘Fiducial, HM UV’ and ‘Fiducial’ having very similar mass growth

histories. We next turn to how adopted galaxy formation physics affects stellar masses using our fiducial simulations suite.

3.1.1 Impact of molecular hydrogen

All of our simulations include H₂ formation, destruction, and cooling (see Section 2.2), with the exception of the test run ‘Fiducial, no H₂ physics’. However, typically we determine star formation based on local density and temperature of all gas (Section 2.3). In one variant (‘Fiducial, H₂ SF’), we tie the star formation exclusively to the molecular component. These two variants thus allow us to determine the overall importance of molecules in our formulation.

In order to reach sufficiently high H₂ gas fractions for star formation, H₂ needs to be self-shielded. In the low-metallicity ISM of the simulated ultra-faint dwarfs (see Section 3.4), this requires high densities, and we find that star formation occurs only at densities $n \sim 10^3$ – 10^4 cm^{-3} in ‘Fiducial, H₂ SF’. As a result, the onset of star formation is slightly delayed ($z \sim 10$) compared to the fiducial model; none the less, as soon as star formation starts, the stellar mass growth quickly catches up with the ‘Fiducial’ galaxy. Overall therefore the star formation law is not strongly sensitive to molecules in itself. This insensitivity may not come as a surprise as star formation anyway is restricted to dense ($n > 300 \text{ cm}^{-3}$) and cold gas ($T < 100$ K), which benefits molecular hydrogen formation (see also Hopkins et al. 2018).

On the other hand, completely neglecting H₂ cooling (‘Fiducial, no H₂ physics’) delays the onset of star formation more significantly (to $z \sim 8$); the final stellar mass is then suppressed by a factor of two relative to ‘Fiducial’. This indicates that the onset of star formation in cold dense gas is sensitive to details of the cooling function. Once star formation starts, enrichment and metal line cooling allows approximately the same gas reservoir to be available for star formation in all models after $z \simeq 8$, limiting the impact

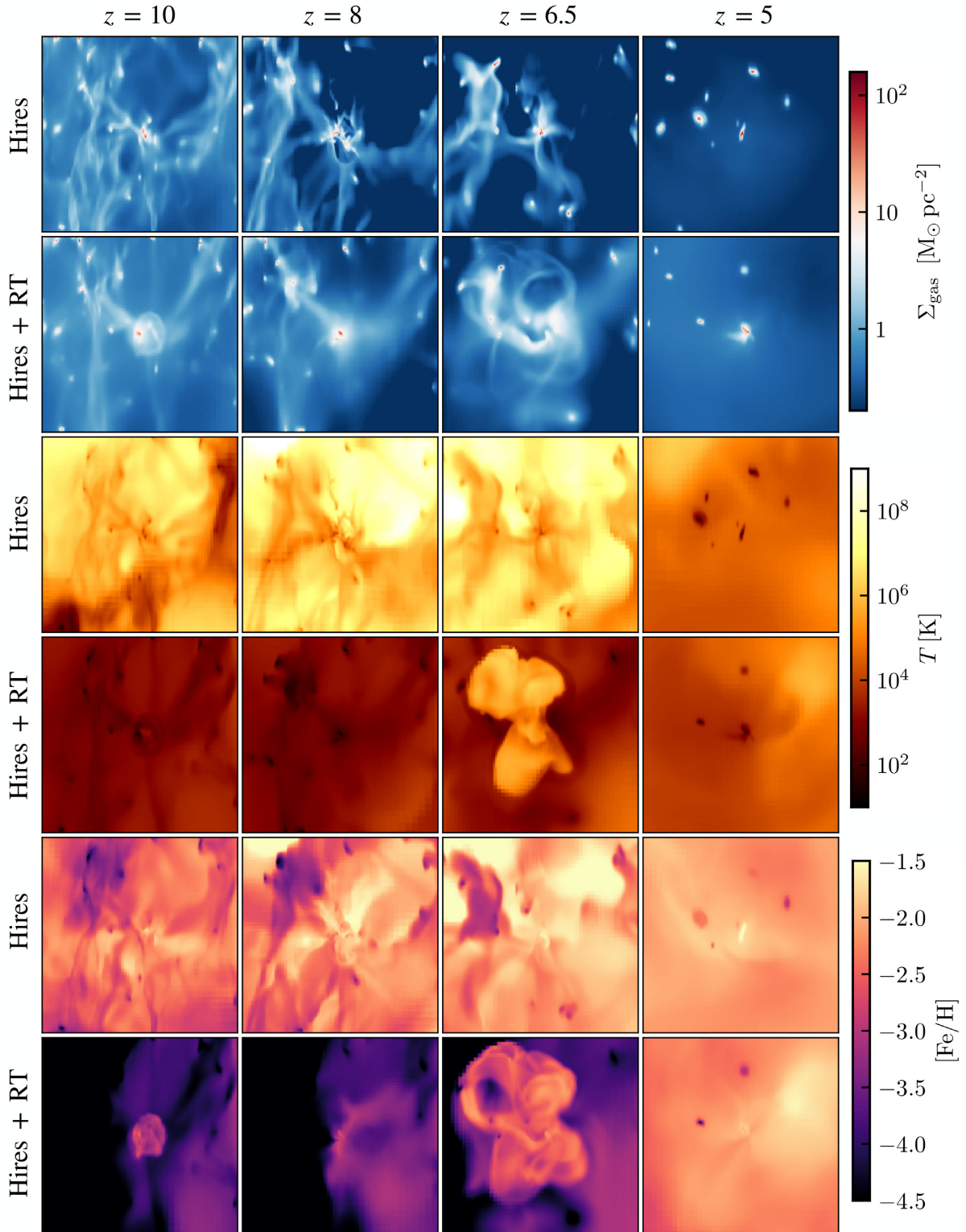


Figure 1. Visual comparison of the gas surface density (top), temperature (middle), and metallicity (bottom) in a $5 \times 5 \text{ kpc}^2$ region (proper kiloparsecs) centred on the main dwarf galaxy in the ‘Hires’ and ‘Hires + RT’ simulations at $z = 10, 8, 6.5$, and 5 . Prior to reionization ($z \gtrsim 6.5$), local radiative feedback in ‘Hires + RT’ completely changes how the galaxy self-regulates; the internal radiation sources are able to keep the gas warm ($T \sim 10^4 \text{ K}$), leading to reduced star formation and weaker outflows at these early times. Conversely, without RT, large-scale SN-driven outflows, together with the significantly weaker contribution from stellar winds, are the only regulation mechanism, resulting in a hot ($T > 10^6 \text{ K}$) and enriched ($[\text{Fe}/\text{H}] > -3$) circumgalactic medium. At late times, the differences in the intergalactic gas in the two cases diminish because cosmic reionization (which is included even in the ‘Hires’ simulation) takes over as the dominant radiation source.

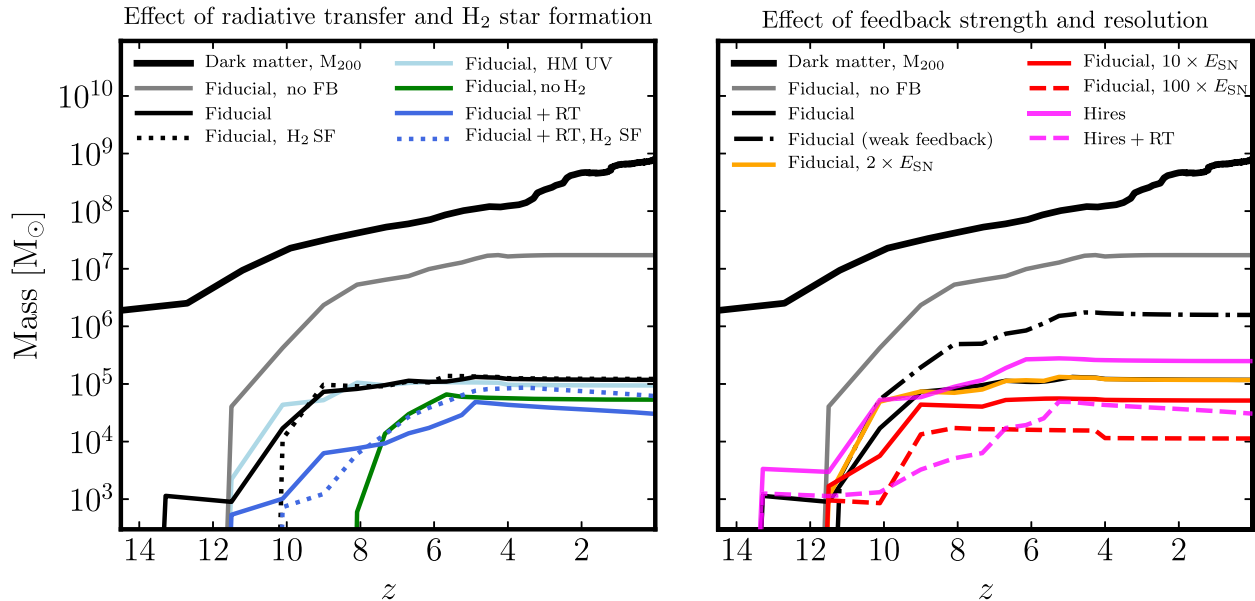


Figure 2. Growth histories of DM and stars from our simulation suite presented in Table 2. Without feedback (‘Fiducial, no FB’), stars form efficiently until $z \sim 4$, with a final stellar mass $M_* > 10^7 M_\odot$. SNe feedback brings the mass to $\sim 10^5 M_\odot$ (‘Fiducial’) or $\sim 2.5 \times 10^5 M_\odot$ (‘Hires’), with radiative feedback suppressing star formation even further ($\sim 3 \times 10^4 M_\odot$, ‘Fiducial + RT’ and ‘Hires + RT’). Changes to other galaxy formation physics, such as the UV background, neglecting H_2 cooling, star formation based on H_2 , can modify the early phases of star formation, but all result in final stellar masses within factor of approximately two to three of each other.

on the final stellar content. We conclude that molecular hydrogen physics is a subdominant uncertainty relative to other factors that we explore below.

3.1.2 Impact of supernova feedback strength

Increasing the available energy per SN explosion by a factor of two (‘Fiducial, $2 \times E_{SN}$ ’) has almost no effect on the stellar mass growth histories. However, an increase by a factor 10 (or 100) results in a suppression by a factor of two (or 10) in final stellar masses. This reveals a weak, sub-linear dependence of the amount of star formation on the strength of SN feedback, compatible with results for simulated disc galaxies (Benincasa et al. 2016).

Limiting the effect of stellar feedback (‘Fiducial, weak feedback’), by artificially imposing numerical ceilings on the allowed SN gas temperature $T_{max} = 10^8$ K and velocities ($v_{fb,max} = 10^3$ km s $^{-1}$), results in an increased stellar mass by an order of magnitude. This stems from the fact that a large fraction of SN explosions occur in low-density gas where SN bubble temperatures and velocities exceed the ceilings (see Section 2.5). Capturing this extreme temperature gas phase, and the associated fast ejecta velocities, is thus essential to capture the full effects of energetic feedback at high resolution.

3.1.3 Impact of radiative feedback

As previously discussed with reference to Fig. 1, including radiative feedback (‘RT’ simulations) generates a major shift in the early behaviour of our galaxy. The ‘Fiducial’ and ‘Hires’ models with RT all result in a final galaxy with stellar masses $M_* \approx 3 \times 10^4 M_\odot$ – an additional factor of ~ 5 –10 reduction from pure SN regulation, and the lowest stellar masses recovered in our simulation suite. Radiative feedback coupled to H_2 -based star formation, i.e. allowing for H_2

destruction by Lyman–Werner radiation from young stars (‘Fiducial + RT, H_2 -based SF’), particularly suppresses star formation rates at early times ($z \gtrsim 6$), although final stellar masses are close to that of ‘Fiducial + RT’.

Without feedback, star formation rates (averaged in age bins of width 100 Myr) reach $\dot{M}_* \sim 10^{-2} M_\odot \text{ yr}^{-1}$ at a lookback time of $t_{\text{lookback}} \sim 12.5$ –13 Gyr. The order of magnitude change in final stellar mass when feedback processes are introduced are mirrored in the average star formation rates, with ‘Fiducial’ and ‘Fiducial + RT’ models peaking at $\sim 10^{-4} M_\odot \text{ yr}^{-1}$ and $\sim \text{few} \times 10^{-5} M_\odot \text{ yr}^{-1}$, respectively, when averaged over 100 Myr windows.

It is worth noting that this effect is far greater than the differences generated by changing numerical resolution. ‘Hires + RT’ forms a total of $3.1 \times 10^4 M_\odot$ compared to 3.0×10^4 for ‘Fiducial + RT’. Enabling RT actually seems to minimize resolution sensitivity, most likely because more gas is kept in a warm and relatively diffuse phase.

3.2 The stellar mass–halo mass relation

In Fig. 3, we show the relation between the galaxy stellar mass and DM halo mass (the M_* – M_{200} relation) for our simulations. For clarity, we show results from the ‘Hires’ models; the corresponding ‘Fiducial’ runs show the same trends, and Table 2 gives data on the entire simulation suite. The left-hand panel shows a comparison to observations, whereas the right-hand panel focuses on comparisons with existing simulation suites. In both cases, we also show the results from abundance matching (AM) as implemented by Read et al. (2017), who used the ‘field’ galaxy stellar mass function from the Sloan Digital Sky Survey (SDSS) and the halo mass function from the cold DM Bolshoi simulation (see also Behroozi, Wechsler & Conroy 2013). Below $M_{200} \sim 5 \times 10^9 M_\odot$

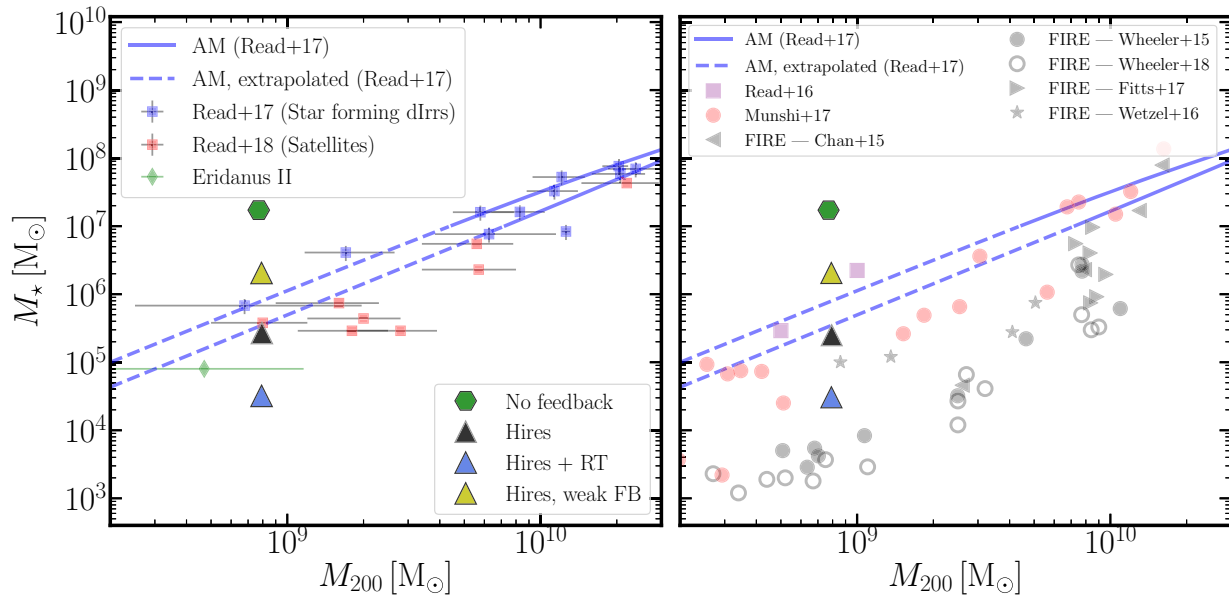


Figure 3. Stellar mass–halo mass relation for our high-resolution simulation suite at $z = 0$, compared (left) to Local Group dwarfs and (right) to existing simulation suites. Abundance matching (AM) extrapolation results from Read et al. (2017) are overplotted on both panels. In the left-hand panel, the blue squares are irregular dwarfs taken from Read et al. (2017) and the red squares are data for (non-star forming) satellite dwarfs compiled by Read & Erkal (2019); the latter are probably the most appropriate point of comparison since our simulated object is always quenched by reionization. A purely SN-driven scenario (‘Hires’) suppresses star formation by two orders of magnitude compared to the model neglecting feedback, and results in a galaxy formation efficiency (M_*/M_{200}) in line with AM. Radiative feedback lowers M_*/M_{200} by almost an additional order of magnitude, within the uncertainties of UFDs such as Eridanus II. In the right-hand panel, the grey points show results from different incarnations of the FIRE project; the red circles give the relation found by Munshi et al. (2017) and the pink squares show the isolated dwarf galaxy simulations of Read et al. (2016a).

($M_* \sim 10^7 M_\odot$), the relation assumes a power-law extrapolation of the SDSS stellar mass function, as indicated by the dashed lines.

Note that the simulated stellar masses are taken ‘as is’, i.e. we do not account for any uncertainties from observational colour fitting procedures. Such uncertainties could lead to observational underestimates of a factor up to approximately two (e.g. Munshi et al. 2013). Moreover, the simulated M_{200} is the mass of DM – we do not include the baryons which at $z = 0$ are negligible.

The left-hand panel data compilation consists of Local Group dwarf galaxies with estimated dynamical masses (Read et al. 2017; Read & Erkal 2019), including galaxies thought to form in halo masses compatible with our simulated galaxies ($M_{200} \sim 10^9 M_\odot$); LeoT, Eridanus-II, and Carina. As discussed by Read & Erkal (2019), the gas-rich star-forming dwarf irregulars have systematically higher estimated galaxy formation efficiencies (M_*/M_{200}) compared to the (non-star-forming) satellites. How some low-mass galaxies manage to be star forming at $z = 0$ is not yet clear; see Wright et al. (2019) for a possible line of explanation. In any case, given that our simulated galaxies are quenched reionization fossils, a direct comparison to the quenched satellites is more appropriate.

The form of the M_{200} – M_* relation in this low-mass regime is highly uncertain, but one can immediately rule out the ‘No feedback’ case as overforming stars by at least one dex. We find that the suppression of stellar mass from SN feedback (‘Hires’, and also in the ‘Fiducial’ case which is not plotted) brings galaxies close to AM predictions and Local Group dwarfs, with $M_*/M_{200} \sim 2.5 \times 10^{-4}$. Almost an additional order of magnitude suppression comes from radiative feedback (‘Hires + RT’),

leading to $M_*/M_{200} \sim 4 \times 10^{-5}$, below the AM extrapolation, but within the uncertainties on Eridanus-II which is the smallest UFD in our observational sample. The model with weak stellar feedback predicts an efficiency of $M_*/M_{200} \sim 2 \times 10^{-3}$ which is in broad agreement with AM and dIrrs – but in tension with quenched UFDs, which are the most appropriate point of comparison.

More observational data on this relationship and its scatter would be helpful to make clearer comparisons to simulations. From the theoretical side, a comparison to existing simulations in the literature is shown in the right-hand panel of Fig. 3, highlighting that different studies currently make very different predictions for the light-to-mass ratio in the faintest objects. We will discuss this point further in Section 4.1.2, but it should already be clear from the figure that there is little agreement on the overall effects of feedback in this regime.

In summary, star formation rates in low-mass dwarf galaxies are extremely sensitive to the detailed feedback physics, with radiative feedback contributing significantly to suppression of star formation. Next, we demonstrate how these differences arise by studying the density and thermal structure of the ISM.

3.3 Structural differences of the interstellar medium

In Fig. 4, we show, for a subset of our simulations in the fiducial suite, the mass-weighted density (left) and the temperature probability distribution functions (PDFs; right) of the gas in the inner parts of the DM halo ($r < 0.25R_{200}$). As the ISM is highly dynamical in low-mass galaxies, a single snapshot in time cannot represent the

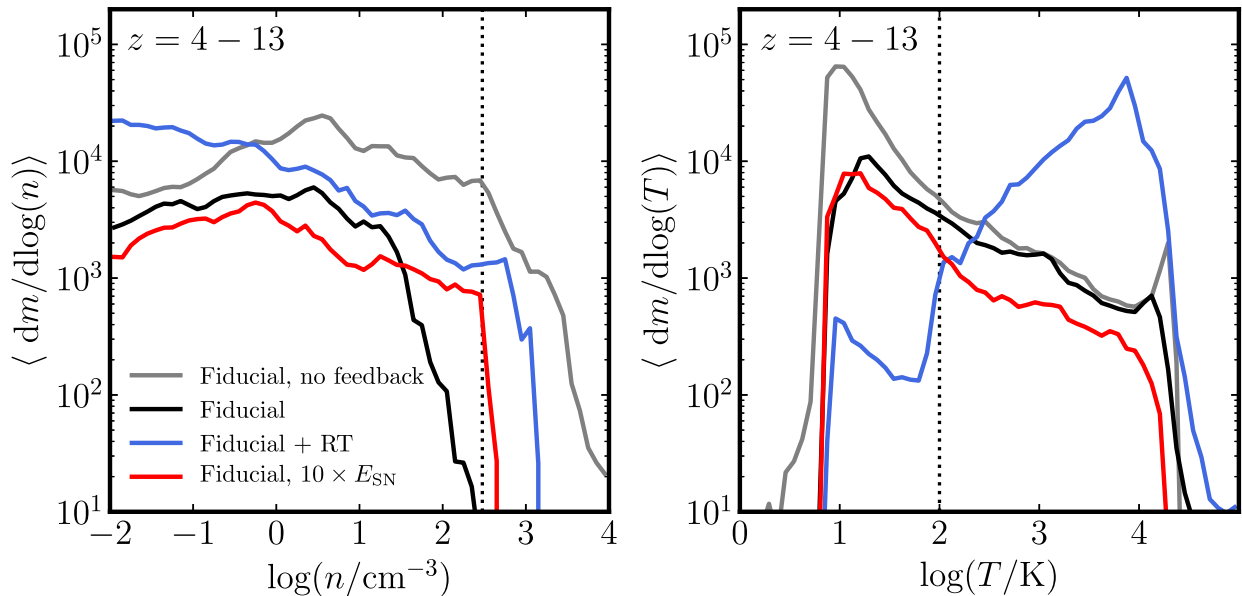


Figure 4. (Left) Mass-weighted density PDFs at $z = 4 - 13$. (Right) Mass-weighted temperature PDFs at $z = 4 - 13$. The grey, black, blue, and red lines indicate no-feedback, ‘Fiducial’, ‘Fiducial + RT’, and ‘Fiducial $10 \times E_{\text{SN}}$ ’ cases. The dotted vertical lines indicate the adopted density ($n = 300 \text{ cm}^{-3}$) and temperature thresholds ($T = 100 \text{ K}$) for star formation. The SN-driven models regulate star formation mainly by blowing out gas from the galaxy, limiting the amount of dense star-forming gas. In contrast, radiative feedback lowers star formation by maintaining most of the ISM in a warm ($T \sim 10^4 \text{ K}$) non-star-forming phase, with less vigorous outflows and more gas remaining in the inner halo at all times.

average state of the galaxies. To alleviate this issue,⁸ the PDFs are mass-weighted averages of all simulation outputs in the redshift range $4 < z < 13$.

Without feedback (grey PDFs), the majority of the ISM’s mass is locked up in a population of dense and cold ($T \approx 10 \text{ K}$) star-forming clouds reaching densities as high as $n \sim 10^4 \text{ cm}^{-3}$. Enabling SN feedback (black PDF) suppresses the fraction of mass in this phase, and reduces the maximum densities by a factor of 100. Note that most gas, by mass, is still retained in the cold gas phase at $T < 100 \text{ K}$ (below our adopted star formation temperature threshold).

Adding radiative feedback (blue PDFs) has a dramatic effect on the temperature structure, with the peak of the temperature PDF shifted from $T \approx 10$ to 10^4 K . The large effect of RT on the thermodynamical state of the ISM, and resulting change in the gas mass available for star formation is the primary mechanism for suppressing galactic star formation rates in these models. The suppression is achieved in a relatively ‘gentle’ fashion, without needing to blow the entire ISM from the galaxy. In fact, we find that before reionization the total gas mass within the galaxy in ‘Fiducial + RT’ is almost identical to the ‘No feedback’ case, despite the two simulations being at opposite extremes in terms of their star formation rates. The total mass in baryons (gas and stars) in the inner halo ($r < 0.25R_{200}$) differ however, with ‘No feedback’ having over three times as much baryonic mass at $z \sim 6$ compared to ‘Fiducial + RT’, and almost 10 times the baryonic mass of ‘Fiducial’.

The gentle suppression of star formation in the ‘Fiducial + RT’ simulation is also evident from the gas density PDF, as it features

more dense gas than the other models, except for ‘Fiducial, no feedback’. Thus star formation is being suppressed without destroying dense clumps. As discussed in Section 1, radiative feedback operates immediately upon the birth of massive stars, which enables the remaining dense cloud to be gently heated. By way of contrast, the first SN in a stellar population only explodes after $\sim 4 \text{ Myr}$, which corresponds to several free-fall times in the dense star-forming gas.⁹ These very different regulation modes will almost certainly have implications for the predicted DM distributions in UFDs which we will study in future work; see also Section 4.

Having established the sensitivity of different sub-grid models and numerical resolution on the star formation rates for this isolated dwarf, in the next section we confront our suite of simulations with observations of nearby dwarf galaxies at $z = 0$.

3.4 Dwarf galaxy scaling relations

In Fig. 5, we compare simulated and observed dwarf galaxy properties. We focus on $z = 0$ relations between V-band magnitudes (M_V), half-mass radii ($r_{1/2}$), stellar velocity dispersions (σ_*), and dynamical mass-to-light ratios (M_{dyn}/L). The results use the same point styles for the four simulations already shown in Fig. 3; the remaining simulations are shown as small circles. The simulations with RT switched on are highlighted in blue; the ‘Fiducial weak feedback’ simulation is highlighted in yellow; the ‘Fiducial $10 \times E_{\text{SN}}$ ’ is shown in red; the ‘Fiducial’ simulation is shown in black; and all other simulations in the suite are shown in grey. All results are also reported in Table 2. Observational data are compiled from

⁸We note that a more robust analysis, not affected by limited time resolution, would have required mass-weighted PDFs to be computed on-the-fly at every simulation time-step.

⁹Note that stellar winds operate before the first SN explosions in all models, but the low metallicity makes them inefficient at regulating star formation.

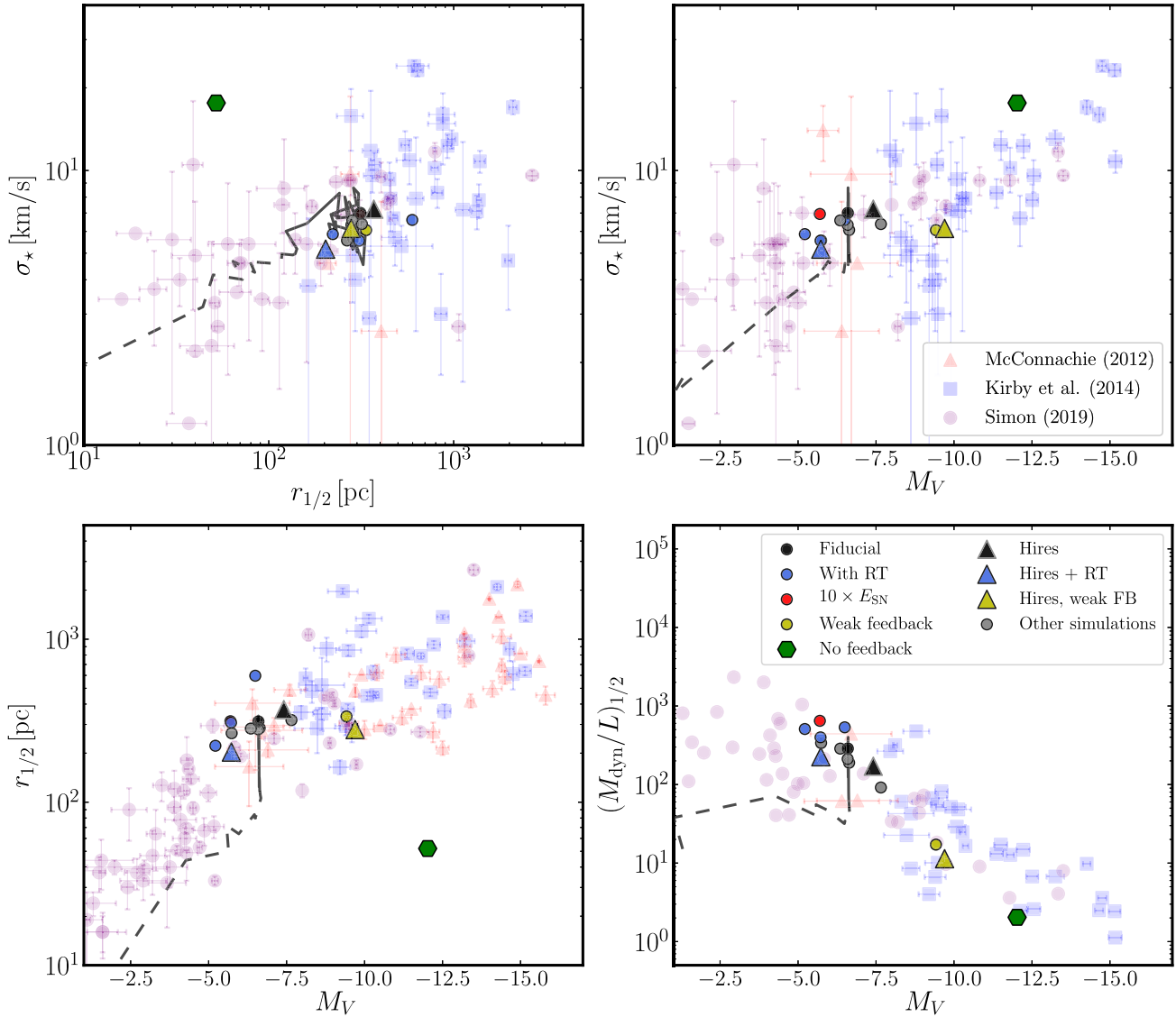


Figure 5. Structural properties of the simulated dwarfs (bold symbols) and the evolutionary track of the ‘Fiducial’ model (with the dashed portion covering the redshift range $4 < z < 13$ prior to quenching). Observational data are taken from McConnell (2012) (red triangles), Kirby et al. (2014) (blue squares), and Simon (2019) (ultra-faints, purple circles). The McConnell and Kirby et al. data both feature MW, M31 and isolated dwarf galaxies, with some overlap. All simulated galaxies, with the exception of the model without feedback, are found to match observations, highlighting that these scaling relations cannot discriminate strongly between different types of galaxy formation physics.

McConnell (2012), Kirby et al. (2013), Kirby et al. (2014), and the recent compilation for UFDs by Simon (2019)¹⁰

V-band magnitudes are computed using all stars in the main dwarf galaxy halo using the SUNSET¹¹ ray-tracing code, assuming a Kroupa IMF (Kroupa 2001). Velocity dispersions are 1D equivalents, i.e. $\sigma_* = \sigma_{*,3D}/\sqrt{3}$, where $\sigma_{*,3D} = \sqrt{\sigma_{*,x}^2 + \sigma_{*,y}^2 + \sigma_{*,z}^2}$. To allow for closer comparison of the dynamical mass-to-light ratios to the observational data, we follow Kirby et al. (2014) and estimate the

dynamical mass using the simple mass estimator:

$$M_{\text{dyn}} = 3\sigma^2 r_{1/2}/G. \quad (3)$$

As can be seen in Fig. 5, all simulations, with exception of the outlier ‘Fiducial, no feedback’, have global galaxy properties compatible with observations. In this set of simulations, M_V lies predominantly in the range -5.5 to -6.5 , with the simulations including RT being the least luminous. Models with ‘weak feedback’ reach $M_V \sim -9.5$. Despite the range in M_V , $z = 0$ galaxy sizes and velocity dispersions are found to be quite insensitive to the different galaxy formation scenarios, with $r_{1/2} \sim$ a few 100 pc and $\sigma_* \sim 5.5$ – 6.5 km s⁻¹ for all galaxies. This weak dependence on final galaxy masses reflects the dominant contribution of the DM halo to the gravitational potential. Indeed, dynamical mass-to-light ratios for all galaxies are $10 < M_{\text{dyn}}/L < 1000$ and are all compatible with observed relations despite the large spread. Again, the model

¹⁰These references contain overlapping sources, sometimes with slightly discrepant values for the quantities under consideration. When data are found to overlap, the latest compilation in Simon (2019) take precedence, and for additional sources the data in Kirby et al. (2013) and Kirby et al. (2014) supersede the data in McConnell (2012).

¹¹Publicly available as a part of the RAMSES distribution.

without feedback is an exception to the above, being very compact ($r_{1/2} = 51$ pc), with a high velocity dispersion ($\sigma_* = 17.6$ km s $^{-1}$) and self-gravitating ($M_{\text{dyn}}/L = 2.0$).

We conclude that it is difficult to rule out – based solely on the observed scaling relations of dwarf galaxies – any of our models in which, at minimum, SN feedback is included. This underlines that, while a large number of studies in the literature have been able to match such relations, they are relatively weak discriminators of the underlying physics governing the faintest dwarfs.

Given the close match to observations, it is interesting to understand whether galaxies evolve ‘along’ observed scaling relations, or if their evolutionary paths are more complex. To this end, Fig. 5 shows evolutionary tracks for the quantities in the representative ‘Fiducial’ model. The dashed lines show $z = 13 - 4$ (i.e. until star formation stops) and the solid lines trace the evolution until $z = 0$. The stellar velocity dispersion shows little evolution, varying by at most a factor of three over cosmic time. In contrast, the half-mass radius evolves significantly, from $r_{1/2} < 10$ pc at $z > 10$ to an over one order of magnitude increase in size at the current epoch. After $z = 4$, most gas is completely removed due to the background UV radiation, and the galaxy expands by factor of three, from $r_{1/2} \sim 100$ pc to $r_{1/2} \sim 300$ pc. Overall, global quantities are found to more or less evolve along the observed $z = 0$ scaling relations, with the possible exception of $r_{1/2}$ and the early evolution ($z > 10$) of the estimated mass-to-light ratio.

3.4.1 The stellar mass–metallicity relation

The final scaling relation we turn to is the stellar mass (or magnitude) versus metallicity relation (MZR). We henceforth refer to the iron to hydrogen abundance ratio ([Fe/H]) as ‘metallicity’. Fig. 6 shows the mean stellar [Fe/H] as a function of M_V (left-hand panel) and M_* (right-hand panel) for all simulations, using the same point styles as in Fig. 5. In the simulations, chemical abundances are calculated for each star particle following (see also Escala et al. 2018),

$$[Y/X] = \log_{10} \left(\frac{f_Y/m_Y}{f_X/m_X} \right) - (\log_{10} \epsilon_{Y,\odot} - \log_{10} \epsilon_{X,\odot}) \quad (4)$$

where Y and X are chemical species, m_Y and m_X are their respective atomic masses, and f_X and f_Y are their respective metal mass fractions. Abundances relative to solar ($\epsilon_{X,\odot}$ and $\epsilon_{Y,\odot}$) are taken from Asplund et al. (2009) (see their table 1). The mean galactic stellar metallicity is computed, motivated by observational measurements of Local Group dwarf galaxies (e.g. Kirby et al. 2013), according to

$$[\text{Fe}/\text{H}] = \frac{\sum_i^N [\text{Fe}/\text{H}]_i m_{*,i}}{\sum_i^N m_{*,i}}, \quad (5)$$

where $[\text{Fe}/\text{H}]_i$ is the metallicity of a star particle, $m_{*,i}$ is the mass of a star particle, and N is the number of star particles in the galaxy.

The left-hand panel contrasts our simulations with data for MW and M31 dwarf spheroidals and Local Group dwarf irregulars taken from Kirby et al. (2013), and UFDs from the recent review by Simon (2019).¹² Observations suggest a metallicity plateau around $[\text{Fe}/\text{H}] \sim -2.5$, and we indicate the lower limit observed in mean $[\text{Fe}/\text{H}]$ (Tucana-II, $[\text{Fe}/\text{H}] = -2.90_{-0.16}^{+0.15}$, Chiti et al. 2018) with a grey dashed line in both panels. This may pose a significant challenge to current numerical simulations which tend to predict near-primordial

abundances for the objects with stellar masses significantly below $10^5 M_\odot$ (right-hand panel), or may indicate that many galaxies at these low luminosities – which are all satellites of the Milky Way and/or M31 – are tidally stripped remnants of once larger systems. We will return to this idea in future work.

By way of contrast, some of our feedback set-ups produce metallicities that are well matched to observations of low-mass dwarfs and UFDs, with $[\text{Fe}/\text{H}] \sim -2.7$ to -2 and $M_V \sim -6.5$ to -5.5 . The MZR is revealed to be a highly sensitive probe of the physics related to the star formation–outflow cycle. For example, boosting the amount of energy injected by SNe leads to strong suppression of $[\text{Fe}/\text{H}]$. Likewise, weaker stellar feedback models lead to an overestimation of $[\text{Fe}/\text{H}]$. This can be understood in terms of our discussion in Section 3.3: *explosive* suppression of star formation using SNe leads to expulsion of enriched gas from the ISM, whereas *continuous* suppression of star formation, using mechanisms like UV heating, enables metals to accumulate within the ISM.

It is also interesting to study the evolutionary tracks in Fig. 6. No version of our galaxy evolves ‘along’ the observed relation, so that star formation quenching at any instance during these formation histories changes the match between theory and observations. Highly explosive suppression of star formation, such as that exhibited by ‘Fiducial, $10 \times E_{\text{SN}}$ ’ and ‘Fiducial, $100 \times E_{\text{SN}}$ ’, results in average metallicities¹³ as low as $[\text{Fe}/\text{H}] \sim -4$. The opposite is true for the models with ‘weak feedback’, which rapidly evolve to almost one dex above the observed relation.

The power of the observed MZR for distinguishing the feedback models in our simulations is encouraging since unlike the scaling relations shown in Fig. 5, it may be possible to apply it as a powerful observational diagnostic of galaxy formation physics in the smallest galaxies. In future work (Orkney et al., in preparation), we will explore this using a larger simulation suite. While SN yields (here taken from Woosley & Weaver 1995), as well as SNIa rates, will introduce uncertainties of factors approximately two at the low metallicities relevant for UFDs (Wiersma et al. 2009), these are likely subdominant compared to the differences between feedback implementations that we find here. A full study will also need to take into account the proximity bias of observed dwarf galaxies with sufficient spectra to derive accurate metallicities (Kirby et al. 2013) and environmental effects such as quenching due to infall on to a larger host galaxy. We discuss this further in Section 4. Finally, we note that the MZR in the faintest regime ($M_V \lesssim -3$) is to date sampled by only a handful of galaxies, and in some cases mean metallicities are poorly determined (for example, Willman 1 has measured metallicities for only two stars). Future deep photometric observations and associated spectroscopic follow-ups will determine the degree to which the MZR is an arbiter of galaxy formation physics.

¹³Allowing for our Pop III metal floor to have solar abundance composition would push up these values to $[\text{Fe}/\text{H}] \sim -3$ (higher than canonical Pop III to Pop II star formation metallicity thresholds, $Z \sim 10^{-4} Z_\odot$, Jaacks et al. 2018), which is closer to, but still in tension with, observations. We note that tests (not shown) with pop III floors at $Z \leq 10^{-4} Z_\odot$ yield close to identical results to those presented here. This means the floor itself, at least in this simulation set-up, can be introduced *a posteriori* to understand features such as the observed plateau in $[\text{Fe}/\text{H}]$.

¹²With UFDs defined, following the approximate magnitude separation by Simon (2019) (see his fig. 5), as galaxies with $M_V > -7.7$.

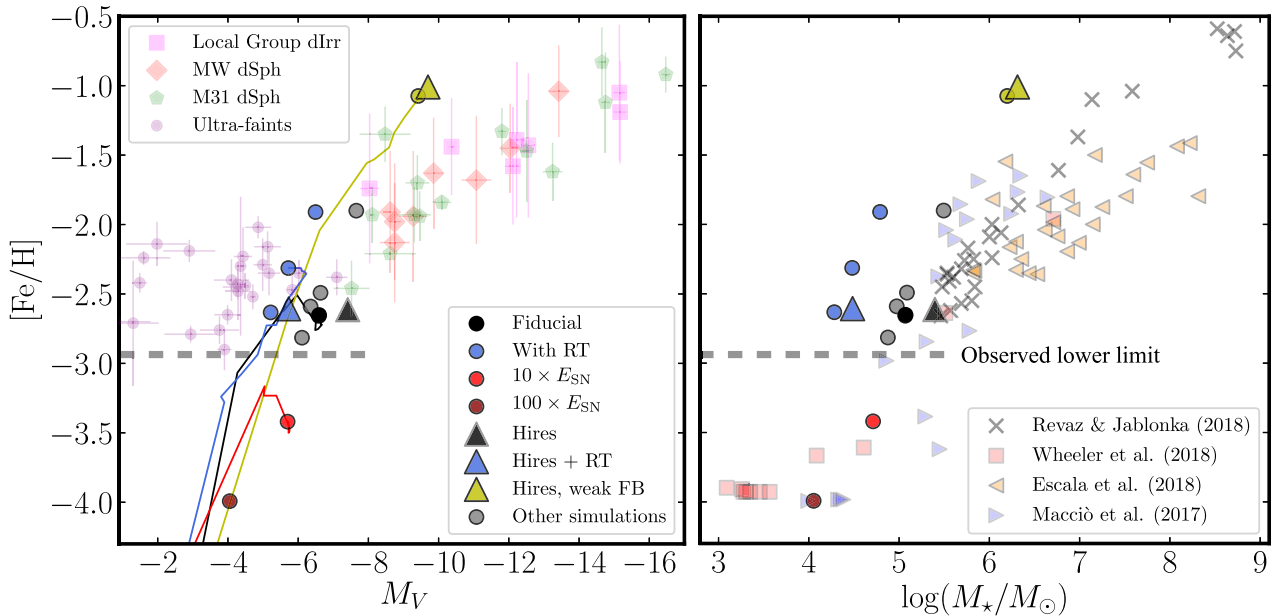


Figure 6. (Left) $[\text{Fe}/\text{H}]$ versus V -band magnitude for the simulated dwarfs (bold symbols), compared with observational data taken from Kirby et al. (2013) (classical Local Group dwarf spheroidal and irregulars: diamonds, pentagons, and squares) and Simon (2019) (ultra-faints: circles), with the error bars showing uncertainties in the mean $[\text{Fe}/\text{H}]$. The grey dashed horizontal line indicates the observed lower limit in mean $[\text{Fe}/\text{H}]$, with the UFD Tucana-II having $[\text{Fe}/\text{H}] = -2.90^{+0.15}_{-0.16}$ (Chiti et al. 2018), the lowest metallicity in our sample. Each of our simulations is plotted at $z = 0$, and the evolution of selected simulations is shown as a line. Models with outflows that are either too strong or too weak fail to match observations, showing that the mass–metallicity relation is a sensitive probe of the feedback mode. (Right) The same metallicities plotted against stellar mass and compared to recent work on zoom simulations of dwarf galaxies in the literature (Macciò et al. 2017; Escala et al. 2018; Revaz & Jablonka 2018), highlighting that many feedback models struggle to reproduce the observed ‘plateau’ of metal abundances in the faintest dwarfs.

4 DISCUSSION

4.1 Comparison with previous work in the literature

In this section, we discuss how our results compare to recent dwarf galaxy formation results in the literature, focusing on dwarf galaxy scaling relations and the M_* – M_{200} relation. Only cosmological simulations reaching a spatial resolution of tens of parsecs are able to resolve the half-mass radii relevant for galaxies forming in haloes of mass $M_{200} \sim 10^8 - 10^{10} M_\odot$ ($r_{1/2} \sim \text{few } 100 \text{ pc}$). As a result, we limit our discussion to zoom simulations (although see Rosdahl et al. 2018, for high-resolution simulations of a uniform volume to $z = 6$).

4.1.1 Effects of feedback on simulated dwarf galaxies

Several authors, using a variety of numerical methods, star formation and feedback prescriptions, report on global properties of simulated dwarfs forming in $M_{200} \sim 10^9 M_\odot$ haloes (including Oñorbe et al. 2015; Wheeler et al. 2015; Fitts et al. 2017; Macciò et al. 2017; Munshi et al. 2017; Revaz & Jablonka 2018). The general picture is that modern simulations produce galaxies compatible with observed scaling relations, with $r_{1/2} \sim 0.1 - 1 \text{ kpc}$, $\sigma_* \sim 5 - 10 \text{ km s}^{-1}$, and stellar masses $M_* \sim 10^4 - 10^6 M_\odot$. All successful models include sub-grid models for stellar feedback that are strong enough to regulate star formation. This keeps dynamical mass-to-light ratios high, allowing stellar velocity dispersions and half-mass radii to be set by the DM halo alone.

While most simulations agree on this basic result, Smith et al. (2019) give an important caveat. Despite the inclusion of a

momentum-capturing SN model similar to ours (see Section 2.3), their simulations all suffer from overcooling, leading to dwarf galaxy properties similar to our ‘Fiducial, No Feedback’ case. This issue was only mitigated when stars in the Smith et al. (2019) simulations were modelled to form at a high star formation efficiency per free-fall time, $\epsilon_{\text{ff}} = 100$ per cent, leading to a more clustered injection of stellar feedback. None the less, while dynamical scaling relations do capture some important aspects of the interaction between feedback algorithms (e.g. ‘blastwave’ versus momentum based feedback) and the clustering of star formation (see e.g. the discussion in Agertz & Kravtsov 2015), we find that metal enrichment is a considerably more sensitive probe of the different feedback models that we have studied in this work.

Wheeler et al. (2015) and Oñorbe et al. (2015) used the FIRE model of feedback (Hopkins et al. 2014, see also the updated FIRE2 model, Hopkins et al. 2018), which treats a range of feedback processes, similar to our model, including a sub-grid model of radiative feedback that includes both radiation pressure and ionizing radiation. The smoothed-particle-hydrodynamics (SPH) simulations by Macciò et al. (2017) and Revaz & Jablonka (2018) both adopted the ‘blastwave’ feedback model (Stinson et al. 2006) that prohibits gas cooling around SNe explosions for an extended period of time, in order to allow for efficient coupling to the ISM. In addition, Macciò et al. (2017) used a phenomenological model for the effect of UV radiation, ‘Early Stellar Feedback’,¹⁴ which is

¹⁴In this, 10 per cent of the UV luminosity of a stellar population – typically 10^{52} erg (10 times the canonical SN explosion energy) – is injected as thermal energy before any SN events take place (Stinson et al. 2013).

likely the reason for them finding lower stellar masses than Revaz & Jablonka (2018) ($M_* \sim 10^4 M_\odot$ compared to $M_* \sim 10^5\text{--}10^6 M_\odot$ for galaxies of mass $M_{200} = 10^9 M_\odot$). The right-hand panel of Fig. 6 shows that the [Fe/H]– M_* relation in Macciò et al. (2017) diverge from the observed one (Kirby et al. 2014) below $M_* \sim 10^6 M_\odot$.

Similar results were recently found by Wheeler et al. (2019) also shown in the right-hand panel of Fig. 6, at a similar numerical resolution to our ‘Hires’ simulations; for $M_* \lesssim 10^5 M_\odot$, their galaxies have mean [Fe/H] < –3.5, with UFDs $M_* \sim 10^3\text{--}10^4 M_\odot$ never enriching at all. This possibly indicates that either star formation is shut down too early in their simulated dwarfs due to overly efficient feedback, or that the IGM around star-forming UFDs, at least in the Local Group, was significantly more pre-enriched by Pop III stars than traditionally thought, as well as predicted by galaxy simulations with Pop III enrichment (e.g. Vandenbroucke, Verbeke & De Rijcke 2016, but see Jaacks et al. 2018), bringing their simulated UFDs closer to the observed [Fe/H] lower limit found for faint satellite galaxies.

4.1.2 The M_* – M_{200} relation

At present, there is no clear theoretical consensus on expectations for M_* for halo masses M_{200} below $\simeq 10^{10} M_\odot$. While individual simulation efforts with fixed star formation and feedback prescriptions tend to produce well-defined M_* – M_{200} relations (e.g. Wheeler et al. 2015; Fitts et al. 2017; Munshi et al. 2017; Wheeler et al. 2019), there are major differences between groups (~ 2 dex, see e.g. Garrison-Kimmel et al. 2017).

The right-hand panel of Fig. 3 collects some examples from the literature. Work using the FIRE feedback and star formation models predicts steeply decreasing galaxy formation efficiencies below $M_{200} \sim 10^{10} M_\odot$, with $M_*/M_{200} < 10^{-5}$ around $M_{200} \sim 10^9 M_\odot$ (close to a few $\times 10^{-6}$, also found in the more recent work by Wheeler et al. 2019). As such, the FIRE simulations predict the lowest galaxy formation efficiencies¹⁵ considered here (together with Macciò et al. 2017), possibly due to strong feedback and the inclusion of ISM heating sources such as photoelectric heating (PEH; see Section 4.2). Our ‘Hires + RT’ model predicts $M_*/M_{200} \sim 5 \times 10^{-5}$, one order of magnitude higher.

The data from the cosmological zoom SPH simulations in Munshi et al. (2017) show a variety of efficiencies, with results in general agreement with our entire simulation suite. By extending their work to include molecular-hydrogen-based star formation, akin to our approach, Munshi et al. (2019) found that stars can only form at densities high enough to allow for gas self-shielding ($n \sim 1000 m_H \text{ cm}^{-3}$), in agreement with our findings for the ‘Fiducial + RT, H₂-based SF’ simulation. This led to a large suppression of the number of galaxies formed below $M_{200} \sim 10^9 M_\odot$. We found essentially no effect when adopting an H₂-based prescription compared to our fiducial approach, which likely stems from us anyway restricting star formation to dense ($n > 300 \text{ cm}^{-3}$) and cold gas ($T < 100 \text{ K}$) that roughly captures the environments where molecular hydrogen formation is possible. Note that we have only studied the effect of H₂ physics for one particular dwarf assembly history, and a larger suite may reveal if H₂-based star formation has an impact. Furthermore, the insensitivity can also stem from us already employing a high-density threshold for star formation in all simulations, or that we are not modelling low enough halo masses, which we leave for a

future investigation. When coupled with RT, the impact of H₂-based star formation led to a ~ 50 per cent difference in the $z = 0$ stellar masses.

There is likely to be a physical spread in M_* at a given M_{200} resulting from the diversity of mass growth histories and the different environments experienced by dwarfs (e.g. Fitts et al. 2017; Garrison-Kimmel et al. 2017; Munshi et al. 2017; Wright et al. 2019). Satellite dwarfs experience DM and stellar mass-loss due to tidal stripping, biasing them to higher M_*/M_{200} ratios (Sawala et al. 2015). However, it should be emphasized that the current scatter in results from the numerical literature is almost certainly dominated by pure feedback discrepancies rather than any systematic differences in the objects being simulated.

In summary, the diversity in M_* – M_{200} at ‘the edge of galaxy formation’ found in the literature reflects in part specific choices in terms of numerical methods, including flavours of sub-grid feedback and star formation prescriptions, as demonstrated by the large scatter found in this paper for a *single* choice of initial condition. In order to interpret observations robustly, both the feedback uncertainty and history variations (see also Revaz & Jablonka 2018) will have to be taken into account; we investigate this in a companion study (Rey et al. 2019).

4.2 Simulation limitations

Having demonstrated that radiative feedback plays a crucial role in regulating the rate of star formation in ultra-faint dwarfs, we now turn to processes we have omitted. Feedback models are still far from complete, but a census of the leading-order missing ingredients and their likely effects helps to interpret the current state of the art.

First, we do not model the effect of PEH, i.e. the heating by far-UV photons from young stars as they liberate electrons from dust grains (Draine 1978). Although this effect is subdominant to UV heating, Forbes et al. (2016) carried out non-cosmological simulations of dwarf galaxies and found PEH to significantly reduce dwarf galaxy stellar masses, more so than SNe feedback. On the other hand, Hu et al. (2017) suggested that the results of Forbes et al. (2016) arose due to incorrect cooling rates in the self-shielded gas. It therefore seems likely that the non-linear coupling between SNe, RT, and dispersal of dense gas, renders PEH subdominant to SN feedback. Furthermore, as the dust-to-gas ratio decreases with metallicity (superlinearly below $0.1 Z_\odot$, Fisher et al. 2014), PEH will most likely be a subdominant effect in the low-metallicity environments studied in this work. None the less, further work in this area is warranted.

Second, we do not include the effect of resonantly scattered Lyman- α photons. Self-consistently modelling the momentum transfer from the scattering of Ly- α photons on to the gas is computationally challenging. Kimm et al. (2018) implemented a sub-grid model of this effect in RAMSES-RT, and argued that in metal-poor systems, Ly- α photons impart momentum comparable to SNe. In their simulations of an idealized disc embedded in $M_{200} = 10^{10} M_\odot$ halo, with SN and RT physics similar to that adopted in this work, star formation rates were reduced by a factor of two. This effect has not yet been studied in a cosmological context. Another missing feedback effect is cosmic rays (CRs). Work by Booth et al. (2013) (see also Hanasz et al. 2013; Salem & Bryan 2014; Chan et al. 2019) demonstrated that the pressure gradients generated by CRs can lead to winds with high mass loading factors, of the order of ~ 10 , in dwarf galaxies. A full treatment CR-driven winds requires magnetohydrodynamics as well as anisotropic CR diffusion coefficients. Furthermore, the impact

¹⁵We note that a few of the Wetzel et al. (2016) ‘Latte’ dwarfs are in line with our results.

of CRs on galactic wind properties depend strongly on the value of the diffusion coefficients, which are only empirically constrained within a factor of 10 for the Milky Way’s ISM, making conclusions on the impact of CRs less robust, at least currently, than e.g. SN feedback.

Finally, each star particle formed in our simulations is assumed to be an SSP with a Kroupa (2001) IMF. In ultra-faint galaxies, the way in which the IMF is populated likely matters, as stochasticity can affect the local level of gas heating from SNe and hence burstiness of star formation (see recent work by Applebaum et al. 2018). Although highly uncertain, if the IMF depends on the local ISM environment by e.g. becoming top-heavy at the low metallicities relevant for UFDs (as suggested by Geha et al. 2013), this can affect galaxy formation in non-trivial ways (Prgomet et al., in preparation). Furthermore, binary star physics can provide additional sources of feedback that have not been modelled here (e.g. Jeon et al. 2015).

We leave the investigation of how the above processes impact galaxy formation in low-mass haloes for future work.

5 CONCLUSIONS

We have carried out cosmological zoom simulations with coupled radiation and hydrodynamics to study the formation of ultra-faint dwarf (UFD) galaxies. We studied galaxy formation in DM haloes, forming in relative isolation, with $z = 0$ virial masses $M_{200} = 10^9 M_{\odot}$, such that the cosmic UV background quenched star formation by $z \sim 4$. The simulations reached a mass and spatial resolution of $\sim 20 M_{\odot}$ and ~ 3 parsecs. Using a single realization of the Gaussian random initial conditions for our dwarf, we investigated the sensitivity of observed galaxy properties to the adopted SN feedback model, UV background, molecular hydrogen (H_2) physics, numerical resolution, and multifrequency RT. Our key findings are as follows:

(i) SN feedback lowers galaxy masses by two orders of magnitude, from $M_{\star} \sim 10^7 M_{\odot}$ when feedback is not included, to $M_{\star} \sim 10^5 M_{\odot}$. Radiative feedback, here self-consistently modelled using multifrequency RT, lowers the galaxy formation efficiency by almost an additional order of magnitude, bringing stellar masses closer to a few $\times 10^4 M_{\odot}$, similar to Local Group ultra-faint dwarfs such as Eridanus-II.

(ii) In models without radiative feedback, we find that dwarf galaxy formation is regulated by vigorous outflows, with much of the ISM being dispersed during starburst episodes. This picture changes when RT is introduced. Radiative feedback acts to keep more gas in a warm ($\sim 10^4$ K) non-star-forming state, leading to less gas accretion and collapse, and, as a result, less vigorous galactic scale outflows, at all times.

(iii) All of our simulations with efficient feedback are in agreement with dynamical measurements of M_{\star} versus M_{200} for Local Group dwarf satellites and UFDs, with $M_{\star}/M_{200} \sim 10^{-4}$. Molecular hydrogen-based star formation and changes to the UV background modifies final stellar masses by at most a factor of two.

(iv) All simulations, even those with artificially weak feedback and $M_{\star}/M_{200} > 10^{-3}$, lie within the scatter of observed scaling relations for V-band magnitudes, half-mass radii, stellar velocity dispersions, and dynamical mass-to-light ratios from Local Group dwarf irregulars, spheroidals, and UFDs, with $r_{1/2} \sim 200\text{--}400$ pc and $\sigma_{\star} \sim 5\text{--}7$ km s $^{-1}$. We find that this insensitivity arises due to structural scaling relations predominantly being set by the host DM halo in galaxies with high mass-to-light-ratios.

(v) We find that the stellar mass–metallicity relation, based on currently available data, can differentiate galaxy formation models, as simulations fail to match observations whenever SN feedback is artificially strong or weak. This highlights that the success of galaxy formation models depends not necessarily on how many observables they can match, but rather on whether they can match a few key observables – in this case the stellar mass–metallicity relation – that are sensitive to changes in the sub-grid physics model.

Our work demonstrates that the enrichment of dwarf galaxies is a more powerful discriminant of feedback processes than any of the other observed scaling relations. The tendency for simulations of ultra-faint dwarfs in the literature to be underenriched may suggest that their feedback models may be excessively violent, unphysically ejecting enriched gas at high redshift, or that the IGM was pre-enriched to much higher levels than traditionally thought. In future work (Rey et al. 2019, Orkney et al., in preparation), we will study a wider range of assembly histories and halo masses in order to understand the generality of this conclusion, as well as to probe how accretion history and environment shape both the metallicity and dynamics of the ultra-faint dwarf population.

ACKNOWLEDGEMENTS

The authors thank the referee Chia-Yu Hu for his constructive comments that improved the quality of the paper. We thank Coral Wheeler, Joshua Simon, Andrew Wetzel, Andrea Macciò, Yvez Revaz, and Alyson Brooks for helpful comments.

OA acknowledges support from the Swedish Research Council (grant 2014-5791) and the Knut and Alice Wallenberg Foundation. AP is supported by the Royal Society. MR acknowledges support from the Perren Fund and the IMPACT fund. JR acknowledges support from the ORAGE project from the Agence Nationale de la Recherche under grant ANR-14-CE33-0016-03.

This work was partially enabled by support from the UCL Cosmoparticle Initiative. This work was performed in part using the DiRAC Data Intensive service at Leicester, operated by the University of Leicester IT Services, which forms part of the **Science and Technology Facilities Council** (STFC) DiRAC HPC Facility (www.dirac.ac.uk). 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. DiRAC is part of the National e-Infrastructure. This work also used the COSMA Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by a BIS National E-infrastructure capital grant ST/K00042X/1, DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. This work was also supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s890. Finally, a large number of simulations for this work were performed on computational resources at LUNARC, the center for scientific and technical computing at Lund University, thanks to financial support from the Royal Physiographic Society of Lund.

REFERENCES

- Abel T., Anninos P., Norman M. L., Zhang Y., 1998, *ApJ*, 508, 518
 Agertz O., Kravtsov A. V., 2015, *ApJ*, 804, 18
 Agertz O., Kravtsov A. V., 2016, *ApJ*, 824, 79
 Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, *ApJ*, 770, 25

- Agertz O., Romeo A. B., Grisdale K., 2015, *MNRAS*, 449, 2156
- Amorisco N. C., 2017, *ApJ*, 844, 64
- Applebaum E., Brooks A. M., Quinn T. R., Christensen C. R., 2018, preprint (arXiv:1811.00022)
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Barkana R., Loeb A., 1999, *ApJ*, 523, 54
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Benincasa S. M., Wadsley J., Couchman H. M. P., Keller B. W., 2016, *MNRAS*, 462, 3053
- Benítez-Llambay A., Frenk C. S., Ludlow A. D., Navarro J. F., 2019, *MNRAS*, 488, 2387
- Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002, *MNRAS*, 333, 177
- Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2846
- Bland-Hawthorn J., Sutherland R., Webster D., 2015, *ApJ*, 807, 154
- Booth C. M., Agertz O., Kravtsov A. V., Gnedin N. Y., 2013, *ApJ*, 777, L16
- Bose S. et al., 2019, *MNRAS*, 486, 4790
- Bovill M. S., Ricotti M., 2009, *ApJ*, 693, 1859
- Bovill M. S., Ricotti M., 2011, *ApJ*, 741, 18
- Brown T. M. et al., 2014, *ApJ*, 796, 91
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bullock J. S., Boylan-Kolchin M., 2017, *ARA&A*, 55, 343
- Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, *ApJ*, 539, 517
- Chan T. K., Keres D., Hopkins P. F., Quataert E., Su K.-Y., Hayward C. C., Faucher-Giguère C.-A., 2019, *MNRAS*, 488, 3716
- Chiti A., Frebel A., Ji A. P., Jerjen H., Kim D., Norris J. E., 2018, *ApJ*, 857, 74
- Christensen C. R., Governato F., Quinn T., Brooks A. M., Shen S., McCleary J., Fisher D. B., Wadsley J., 2014, *MNRAS*, 440, 2843
- Contenta F. et al., 2018, *MNRAS*, 476, 3124
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- Dalla Vecchia C., Schaye J., 2008, *MNRAS*, 387, 1431
- Decataldo D., Pallottini A., Ferrara A., Vallini L., Gallerani S., 2019, *MNRAS*, 487, 3377
- Dekel A., Silk J., 1986, *ApJ*, 303, 39
- Draine B. T., 1978, *ApJS*, 36, 595
- Dutton A. A., Macciò A. V., Buck T., Dixon K. L., Blank M., Obreja A., 2019, *MNRAS*, 486, 655
- Efstathiou G., 1992, *MNRAS*, 256, 43P
- Efstathiou G., 2000, *MNRAS*, 317, 697
- Eisenstein D. J., Hut P., 1998, *ApJ*, 498, 137
- Escala I. et al., 2018, *MNRAS*, 474, 2194
- Fall S. M., Efstathiou G., 1980, *MNRAS*, 193, 189
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761
- Fisher D. B. et al., 2014, *Nature*, 505, 186
- Fitts A. et al., 2017, *MNRAS*, 471, 3547
- Forbes J. C., Krumholz M. R., Goldbaum N. J., Dekel A., 2016, *Nature*, 535, 523
- Garrison-Kimmel S., Bullock J. S., Boylan-Kolchin M., Bardwell E., 2017, *MNRAS*, 464, 3108
- Gatto A., Fraternali F., Read J. I., Marinacci F., Lux H., Walch S., 2013, *MNRAS*, 433, 2749
- Geha M. et al., 2013, *ApJ*, 771, 29
- Glover S. C. O., Clark P. C., 2012, *MNRAS*, 421, 9
- Gnedin N. Y., Kaurov A. A., 2014, *ApJ*, 793, 30
- Gnedin N. Y., Kravtsov A. V., 2006, *ApJ*, 645, 1054
- Gnedin N. Y., Kravtsov A. V., 2010, *ApJ*, 714, 287
- Gnedin N. Y., Tassis K., Kravtsov A. V., 2009, *ApJ*, 697, 55
- Grebel E. K., Gallagher J. S., III, Harbeck D., 2003, *AJ*, 125, 1926
- Grisdale K., Agertz O., Romeo A. B., Renaud F., Read J. I., 2017, *MNRAS*, 466, 1093
- Grisdale K., Agertz O., Renaud F., Romeo A. B., 2018, *MNRAS*, 479, 3167
- Grisdale K., Agertz O., Renaud F., Romeo A. B., Devriendt J., Slyz A., 2019, *MNRAS*, 486, 5482
- Guillet T., Teyssier R., 2011, *J. Comput. Phys.*, 230, 4756
- Haardt F., Madau P., 1996, *ApJ*, 461, 20
- Hanasz M., Lesch H., Naab T., Gawryszczak A., Kowalik K., Wóltański D., 2013, *ApJ*, 777, L38
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
- Hopkins P. F. et al., 2018, *MNRAS*, 480, 800
- Hu C.-Y., Naab T., Glover S. C. O., Walch S., Clark P. C., 2017, *MNRAS*, 471, 2151
- Iwamoto N., Umeda H., Tominaga N., Nomoto K., Maeda K., 2005, *Science*, 309, 451
- Jaacks J., Thompson R., Finkelstein S. L., Bromm V., 2018, *MNRAS*, 475, 4396
- Jeon M., Bromm V., Pawlik A. H., Milosavljević M., 2015, *MNRAS*, 452, 1152
- Jethwa P., Erkal D., Belokurov V., 2018, *MNRAS*, 473, 2060
- Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19
- Kim C.-G., Ostriker E. C., 2015, *ApJ*, 802, 99
- Kimm T., Katz H., Haehnelt M., Rosdahl J., Devriendt J., Slyz A., 2017, *MNRAS*, 466, 4826
- Kimm T., Haehnelt M., Blaizot J., Katz H., Michel-Dansac L., Garel T., Rosdahl J., Teyssier R., 2018, *MNRAS*, 475, 4617
- Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, *ApJ*, 779, 102
- Kirby E. N., Bullock J. S., Boylan-Kolchin M., Kaplinghat M., Cohen J. G., 2014, *MNRAS*, 439, 1015
- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Kravtsov A. V., 2003, *ApJ*, 590, L1
- Kravtsov A. V., Gnedin O. Y., Klypin A. A., 2004, *ApJ*, 609, 482
- Kroupa P., 2001, *MNRAS*, 322, 231
- Krumholz M. R., 2012, *ApJ*, 759, 9
- Krumholz M. R., Gnedin N. Y., 2011, *ApJ*, 729, 36
- Krumholz M. R., Tan J. C., 2007, *ApJ*, 654, 304
- Lee E. J., Miville-Deschênes M.-A., Murray N. W., 2016, *ApJ*, 833, 229
- Leitherer C. et al., 1999, *ApJS*, 123, 3
- Leitner S. N., 2012, *ApJ*, 745, 149
- Levermore C. D., 1984, *J. Quant. Spectrosc. Radiat. Transfer*, 31, 149
- Macciò A. V., Frings J., Buck T., Penzo C., Dutton A. A., Blank M., Obreja A., 2017, *MNRAS*, 472, 2356
- Martizzi D., Faucher-Giguère C.-A., Quataert E., 2015, *MNRAS*, 450, 504
- Mashchenko S., Wadsley J., Couchman H. M. P., 2008, *Science*, 319, 174
- Mayer L., Mastropietro C., Wadsley J., Stadel J., Moore B., 2006, *MNRAS*, 369, 1021
- McConnachie A. W., 2012, *AJ*, 144, 4
- Munshi F. et al., 2013, *ApJ*, 766, 56
- Munshi F., Brooks A. M., Applebaum E., Weisz D. R., Governato F., Quinn T. R., 2017, preprint (arXiv:1705.06286)
- Munshi F., Brooks A. M., Christensen C., Applebaum E., Holley-Bockelmann K., Quinn T. R., Wadsley J., 2019, *ApJ*, 874, 40
- Murray N., 2011, *ApJ*, 729, 133
- Nickerson S., Teyssier R., Rosdahl J., 2018, *MNRAS*, 479, 3206
- Ocvirk P. et al., 2018, preprint (arXiv:1811.11192)
- Oñorbe J., Boylan-Kolchin M., Bullock J. S., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., Quataert E., Murray N., 2015, *MNRAS*, 454, 2092
- Planck Collaboration XVI, 2014, *A&A*, 571, A16
- Pontzen A., Governato F., 2012, *MNRAS*, 421, 3464
- Read J. I., Erkal D., 2019, *MNRAS*, 487, 5799
- Read J. I., Gilmore G., 2005, *MNRAS*, 356, 107
- Read J. I., Wilkinson M. I., Evans N. W., Gilmore G., Kleyna J. T., 2006a, *MNRAS*, 367, 387
- Read J. I., Pontzen A. P., Viel M., 2006b, *MNRAS*, 371, 885
- Read J. I., Agertz O., Collins M. L. M., 2016a, *MNRAS*, 459, 2573
- Read J. I., Iorio G., Agertz O., Fraternali F., 2016b, *MNRAS*, 462, 3628
- Read J. I., Iorio G., Agertz O., Fraternali F., 2017, *MNRAS*, 467, 2019
- Revaz Y., Jablonka P., 2018, *A&A*, 616, A96
- Rey M. P., Pontzen A., 2018, *MNRAS*, 474, 45

- Rey M. P., Pontzen A., Agertz O., Orkney M. D. A., Read J. I., Saintonge A., Pedersen C., 2019, *ApJL*, 886, L3
- Rosdahl J., Blaizot J., 2012, *MNRAS*, 423, 344
- Rosdahl J., Teyssier R., 2015, *MNRAS*, 449, 4380
- Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, *MNRAS*, 436, 2188
- Rosdahl J., Schaye J., Teyssier R., Agertz O., 2015, *MNRAS*, 451, 34
- Rosdahl J. et al., 2018, *MNRAS*, 479, 994
- Rosen A., Bregman J. N., 1995, *ApJ*, 440, 634
- Roth N., Pontzen A., Peiris H. V., 2016, *MNRAS*, 455, 974
- Saitoh T. R., Daisaka H., Kokubo E., Makino J., Okamoto T., Tomisaka K., Wada K., Yoshida N., 2008, *PASJ*, 60, 667
- Salem M., Bryan G. L., 2014, *MNRAS*, 437, 3312
- Sanders J. L., Evans N. W., Dehnen W., 2018, *MNRAS*, 478, 3879
- Sawala T. et al., 2015, *MNRAS*, 448, 2941
- Simon J. D., 2019, *ARA&A*, 57, 375
- Smith M. C., Sijacki D., Shen S., 2019, *MNRAS*, 485, 3317
- Springel V., Frenk C. S., White S. D. M., 2006, *Nature*, 440, 1137
- Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, *MNRAS*, 373, 1074
- Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, *MNRAS*, 428, 129
- Teyssier R., 2002, *A&A*, 385, 337
- Tollet E. et al., 2016, *MNRAS*, 456, 3542
- Toro E. F., Spruce M., Speares W., 1994, *Shock Waves*, 4, 25
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, 438, 1985
- Vandenbroucke B., Verbeke R., De Rijcke S., 2016, *MNRAS*, 458, 912
- Verbeke R., Vandenbroucke B., Rijcke S. D., 2015, *ApJ*, 815, 85
- Vikhlinin A. et al., 2009, *ApJ*, 692, 1060
- Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, *ApJ*, 789, 148
- Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, *ApJ*, 827, L23
- Wheeler C. et al., 2019, *MNRAS*, 490, 4447
- Wheeler C., Oñorbe J., Bullock J. S., Boylan-Kolchin M., Elbert O. D., Garrison-Kimmel S., Hopkins P. F., Kereš D., 2015, *MNRAS*, 453, 1305
- White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341
- Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, *MNRAS*, 399, 574
- Wise J. H., Abel T., Turk M. J., Norman M. L., Smith B. D., 2012a, *MNRAS*, 427, 311
- Wise J. H., Turk M. J., Norman M. L., Abel T., 2012b, *ApJ*, 745, 50
- Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, *MNRAS*, 442, 2560
- Woosley S. E., Weaver T. A., 1995, *ApJS*, 101, 181
- Wright A. C., Brooks A. M., Weisz D. R., Christensen C. R., 2019, *MNRAS*, 482, 1176

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.