Astronomy & Astrophysics

How much metal did the first stars provide to the ultra-faint dwarfs?

Mahsa Sanati[1] , Fabien Jeanquartier[1], Yves Revaz[1], and Pascale Jablonka[1,2]

[1] Institute of Physics, Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), 1290 Sauverny, Switzerland
   e-mail: mahsa.sanati@epfl.ch
[2] GEPI, CNRS UMR 8111, Observatoire de Paris, PSL University, 92125 Meudon Cedex, France

ABSTRACT

Numerical simulations of dwarf galaxies have so far failed to reproduce the observed metallicity-luminosity relation, down to the regime of ultra-faint dwarfs (UFDs). We address this issue by exploring how the first generations of metal-free stars (Pop III) could help increase the mean metallicity ([Fe/H]) of those small and faint galaxies. We ran zoom-in chemo-dynamical simulations of 19 halos extracted from a Λ Cold Dark Matter (CDM) cosmological box and followed their evolution down to redshift $z = 0$. Models were validated not only on the basis of galaxy global properties, but also on the detailed investigation of the stellar abundance ratios ($[\alpha/\text{Fe}]$). We identified the necessary conditions for the formation of the first stars in mini-halos and derived constraints on the metal ejection schemes. The impact of Pop III stars on the final metallicity of UFDs was evaluated by considering different stellar mass ranges for their initial mass function (IMF), the influence of pair-instability supernovae (PISNe), and their energetic feedback, as well as the metallicity threshold that marks the transition from the first massive stars to the formation of low-mass long-lived stars. The inclusion of Pop III stars with masses below $140\,M_\odot$, and a standard IMF slope of -1.3 does increase the global metallicity of UFDs, although these are insufficient to resolve the tension with observations. The PISNe with progenitor masses above $140\,M_\odot$ do allow the metal content of UFDs to further increase. However, as PISNe are very rare and sometimes absent in the faintest UFDs, they have a limited impact on the global faint end of the metallicity-luminosity relation. Despite a limited number of spectroscopically confirmed members in UFDs, which make the stellar metallicity distribution of some UFDs uncertain, our analysis reveals that this is essentially the metal-rich tail that is missing in the models. The remaining challenges are thus both observational and numerical: (i) to extend high-resolution spectroscopy data samples and confirm the mean metallicity of the faintest UFDs; and (ii) to explain the presence of chemically enriched stars in galaxies with very short star formation histories.

Key words. galaxies: dwarf – stars: Population III – Galaxy: abundances – methods: numerical

1. Introduction

Ultra-faint dwarf galaxies (UFDs) are the faintest galaxies known, with V-band luminosities fainter than $L_V = 10^5\,L_\odot$, $M_V < -7.7$ (see Simon 2019, for a recent review). Some of them are as faint as a few hundred solar luminosities and may be as compact as faint globular clusters. Ultra-faint dwarfs are also the most dark-matter-dominated galactic systems and, as such, they constitute fundamental probes of the cosmological model (Bullock & Boylan-Kolchin 2017).

Semi-analytical and hydro-dynamical numerical simulations have shown that UFDs are compatible with the first galaxies formed in mini-halos before the epoch of reionization (e.g., Ricotti & Gnedin 2005; Wyithe & Loeb 2006; Salvadori & Ferrara 2009; Bovill & Ricotti 2009; Wheeler et al. 2019; Rodriguez Wimberly et al. 2019). These metal poor systems (Muñoz et al. 2006; Martin et al. 2007; Simon & Geha 2007; Kirby et al. 2008) had indeed formed the bulk of their stellar mass (≥80%) by $z \sim 6$, as revealed by deep ground- and space-based color-magnitude diagrams (Okamoto et al. 2012; Brown et al. 2014; Sacchi et al. 2021; Gallart et al. 2021). Therefore, their properties offer a unique insight into the physics at work in the early Universe.

One might think that these small systems, with very short star formation histories, would be the product of simple physical processes. This is not the case. They do challenge numerical simulations, which struggle to reproduce their properties. Their small size, low mass, and shallow gravitational potential makes them extremely sensitive to the numerical implementation of the physical processes; they require high resolution, and they are critically sensitive to a correct and well controlled treatment of the first generations of stars, whose chemical signatures should be more salient than in more massive systems (e.g., Ji et al. 2015).

Along this line, one of the most striking obstacles faced by numerical simulations is the reproduction of the observed luminosity-metallicity relation in UFDs. As described in detail in Sect. 2, all hydro-dynamical models systematically predict too low a stellar metallicity at a given luminosity, compared to the observations, or in other words, too steep a slope of the [Fe/H]-L_V relation (Jeon et al. 2017, 2021; Macciò et al. 2017; Escala et al. 2018; Wheeler et al. 2019; Agertz et al. 2020; Applebaum et al. 2021; Prgomet et al. 2022). The question arises whether the models are so far missing an important source of metals, and/or whether the nucleosynthesis products are too diluted in the interstellar medium (ISM) after the supernovae explosions.

The pair-instability supernovae (PISNe), first introduced by Barkat et al. (1967), are, in theory, an outstanding source of metals, and hence worth considering in the context of UFDs. These supernovae result from an instability induced by the production of free electron-positron pairs in metal-free stars with masses larger than $140\,M_\odot$. This instability leads to a complete disruption of the star, releasing a large amount of heavy elements. Since UFDs cease forming stars shortly after cosmic

reionization, PISNe could significantly influence their chemical evolution as part of the first (population III, hereafter Pop III) stars, pre-enriching the ISM. The influence of Pop III stars in UFDs has hardly been investigated so far. They are either included in the semi-analytical approach of Salvadori & Ferrara (2009), or in the hydro-dynamical simulations of Jeon et al. (2017), where their specific influence has not been documented.

No metal-free stars have yet been observed, suggesting that Pop III stars were sufficiently massive to either collapse into black holes or to explode as a supernovae. In the latter case, they could leave their nucleosynthesis imprints in the next generations of low-mass long-lived stars (population II, hereafter Pop II) that are observed today. However, the physical conditions under which they are formed and their detailed properties are still unsettled. This leaves wide open questions about the time frame of their formation in the cosmic history (Bromm & Larson 2004; Yoshida et al. 2012; Hummel et al. 2012; Magg et al. 2016), the mass of the halos in which they were hosted, and their external enrichment from neighboring halos (see e.g., Skinner & Wise 2020; Schauer et al. 2021; Graziani et al. 2017; Hicks et al. 2021), as well as their initial mass function (IMF, see e.g., Machida 2008; Turk et al. 2009; Stacy et al. 2010; Clark et al. 2011; Maeder & Meynet 2012; de Bennassuti et al. 2017; Chon et al. 2021, 2022). Chemo-dynamical cosmological zoom-in simulations of UFDs can test different Pop III models and verify their compatibility with the stellar abundance ratios observed in Pop II stars.

The goal of this paper is to precisely track down the conditions under which the global metal content of the model UFDs, traced by their iron abundance, can be increased at fixed stellar mass and/or luminosity. This involves investigating: (i) the physical conditions of star formation; (ii) the possible strong Fe contributors among the very first stars; (iii) the IMF of the first stars; (iv) how metals are released and mixed in the ISM; and (v) the impact of the supernovae feedback energy. Just as in our previous works, we discriminate between different scenarios not only on the grounds of global galaxy properties (e.g., $\langle[Fe/H]\rangle$, luminosities, and mass) but also on their detailed stellar abundance ratios, stellar metallicity distributions, and velocity dispersions.

The paper is organized as follows. In Sect. 2, we review in detail the discrepancies between the observed and simulated luminosity-metallicity relations, and the different scenarios that have been suggested to address this issue. In Sect. 3, we present our numerical methods, introduce in particular our treatment of the Pop III stars, and describe the entire set of simulations. In Sect. 4, we discuss the conditions from the formation of high-mass and low-mass stars, the IMF of Pop III stars, and their explosion energy, as well as their impact on the relation between stellar metallicity and luminosity, with and without PISNe. We also compare the α-elements abundance ratios, as well as the metallicity distribution function in models and observations. Our conclusions are presented in Sect. 5.

2. The challenge of the luminosity-metallicity relation

Similarly to more massive galaxies, dwarf galaxies fall on the scaling relations. Among those, the metallicity-luminosity relation was first observed by Lequeux et al. (1979) for irregular and blue compact galaxies, and subsequently confirmed by many authors, extending both to dwarf galaxies or more distant spirals (see e.g., Skillman et al. 1989; Garnett 2002; Tremonti et al. 2004). Basically, dwarf galaxies brighter than $L_V > 10^5\,L_\odot$ synthesize a larger quantity of metals, which are further locked in
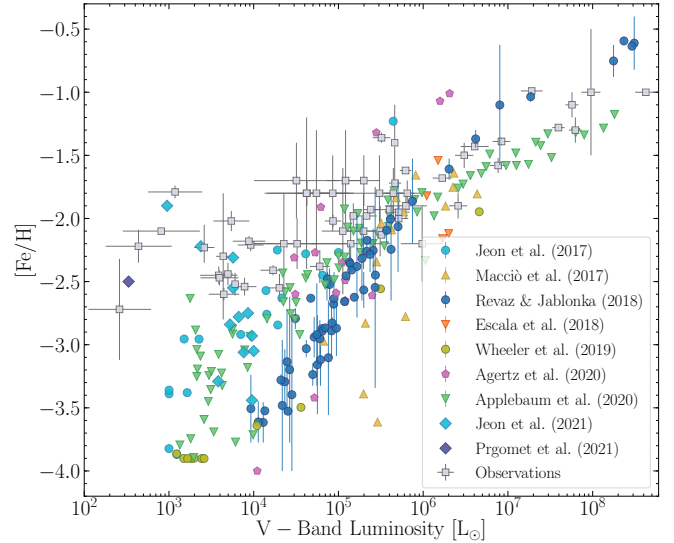

Fig. 1. Comparison of the luminosity-metallicity relation for dwarf galaxies and UFDs, between Local Group observations and simulations. The gray squares represent our Local Group sample (see text for details). Colored points stand for different simulations published with the references given at the bottom of the figure. We extended our own work (Revaz & Jablonka 2018) with the results of unpublished galaxies in the UFD regime. Error bars for the stellar metallicity in our simulated galaxies were calculated by comparing the results obtained with the mean, the median, and the mode acquired with several binning choices. See Sect. 3.6 for more details on the method used.

the next generation of stars. This is relatively well reproduced by the models. However, in the faint (UFD) regime, the observed dispersion in stellar metallicity at fixed luminosity increases (Simon 2019), and the slope of the observed relation is less steep than the models, with too low a [Fe/H] at a given luminosity. This common feature remains despite a variety of numerical schemes and stellar feedback recipes (Agertz et al. 2020). While this systematic deviation drew attention only recently, in particular with the recent improvements in numerical resolution, it was already mentioned in past studies (Bovill & Ricotti 2009, see their Fig. 3).

This discrepancy is illustrated in Fig. 1, where Local Group dwarfs and UFDs are compared to the available hydro-simulations from Jeon et al. (2017), Macciò et al. (2017), Revaz & Jablonka (2018), Escala et al. (2018), Wheeler et al. (2019), Agertz et al. (2020), Applebaum et al. (2021), and Prgomet et al. (2022). This figure includes a set of UFDs not published in Revaz & Jablonka (2018) but obtained with exactly the same physical model. The observed galaxy sample was obtained thanks to the continuously updated Local Group and Nearby Dwarf Galaxies database[1] of McConnachie (2012) and are represented by gray squares. It should be noted that our sample includes only confirmed UFDs that benefit from medium resolution spectroscopy with metallicity derived either from spectral synthesis or Calcium triplet (CaT) calibration.

Several explanations are proposed for the origin of the tension between models and observations. Bovill & Ricotti (2011) suggest that the observed dwarfs may be a subset of bright primordial fossils that have been stripped of 90%–99% of their stars. Indeed, tidal striping could remove stars from a dwarf galaxy, subsequently reducing its luminosity without affecting

---

[1] http://www.astro.uvic.ca/~alan/Nearby_Dwarf_Database.html

its metallicity. However, only dwarf galaxies with a a very small pericenter will lose an appreciable fraction of their stars (Peñarrubia et al. 2008). Moreover, striping would induce an expansion of the stellar component, in contradiction with UFDs having a very small half-light radius. Environmental pre-enrichment by the host galaxy was suggested by Wheeler et al. (2019). Applebaum et al. (2021) mention, however that in a Milky Way-like environment, dwarf galaxies that have never been harassed or stripped show comparable metallicities with respect to the dwarf satellites. They suggest instead that their simulated dwarfs are simply under-producing iron and that the tension can be solved by increasing the contribution to the iron abundance of Type Ia supernovae (SNeIa). This, however, requires either reducing the SNeIa explosion timescale, or extending the star formation history of UFDs beyond two Gyr, in contradiction with current observations.

Prgomet et al. (2022) looked into the impact of a metallicity-dependent IMF on the stellar mass and metallicity of their simulated dwarf galaxies. They examined the effect of a top-heavy IMF for metal-poor stars. They show that the largest number of massive stars exploding as core-collapse supernovae (CCSNe), decreases the final luminosity of the system through enhanced stellar feedback, while it increases the metallicity. Unfortunately, the boosted SNe feedback also leads to highly diffuse systems with a half-light radius of ~0.5 dex larger than the observed values.

Gandhi et al. (2022) suggest boosting the iron production by using a metallicity-dependent SNeIa rate. However, because the number of SNeIa is low in UFDs, the resulting model does not exceed [Fe/H] $\cong$ −3, which is still below the observed metallicities.

Alternatively, Revaz & Jablonka (2018) suggest that Pop III stars could play a role in reconciling observations with simulations, a feature worth exploring in hydro-codes. Indeed, hydro-codes self-consistently treat the formation of galaxies, in particular the merger history of sub-halos that may have been enriched independently. Moreover, they properly account for the ejecta of polluted gas and its recycling after re-accretion on the same halo it originates, or on different ones.

3. Methods and simulations

We performed a set of 19 zoom-in Λ Cold Dark Matter (CDM) cosmological simulations of dwarf and UFD galaxies with a final luminosity between 2×10^3 to 3×10^5 $L_\odot$. We selected galaxy halos from different regions in the dark-matter-only $(5\,\mathrm{Mpc})^3$ cosmological volume used in Revaz & Jablonka (2018). Those regions were identified to form a dark halo with a mass typical of a UFD at redshift $z = 0$. We re-simulated those halos using a zoom-in technique including a full treatment of baryons that we describe below. The initial conditions were obtained with the MUSIC code (Hahn & Abel 2011). Compared to our previous studies of dwarf galaxies (e.g., Revaz & Jablonka 2018; Harvey et al. 2018; Hausammann et al. 2019; Sanati et al. 2020), simulations in this work were run with a factor of eight enhancement in resolution for the gas and dark matter, and a factor of 16 for the stars. This was essential to resolve the Pop III host halos with masses between 10^6 and 10^7 $M_\odot$, and to guarantee that UFDs with a luminosity as low as a few 10^3 $L_\odot$ were populated with at least ten stellar particles. In the refined region (the so-called Lagrangian region), we achieved a dark matter, gas, and stellar mass resolution of 4172 $M_\odot$, 761 $M_\odot$, and 380 $M_\odot$, respectively. The resolution was gradually degraded from level 10 to

6 outside[2]. The initial Lagrangian region was an ellipsoidal of size of about 1.3 Mpc across, to enclose all the particles that eventually ended up within the target halo at redshift $z = 0$ (Oñorbe et al. 2014a). It is important to note that we limited our stellar resolution to 380 $M_\odot$ to ensure a proper IMF sampling up to the most massive stars in the mass range [140−300] $M_\odot$ considered in this work (see Sects. 3.3 and 3.2).

We used the cosmology of the Planck Collaboration XIII (2016) with $\Omega_\Lambda = 0.685$, $\Omega_m = 0.315$, $\Omega_b = 0.0486$, $H_0 = 67.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $n_s = 0.9603$, and $\sigma_8 = 0.829$. All simulations were started at redshift $z = 70$, ensuring that the rms variance of the initial density field, σ_8, lied between 0.1 and 0.2 (Knebe et al. 2009; Oñorbe et al. 2014a), and ran till redshift $z = 0$.

All galaxies in this work were evolved in a cosmological context but without taking into account any interaction with a Milky Way-like system. Such an interaction might potentially influence the star formation history of UFDs. However, there is no clear observational evidence that the star formation histories of the Local Group satellites depend on such interactions with the Milky Way (Dixon et al. 2018).

3.1. Numerical setup

The simulations were run with GEAR (Revaz & Jablonka 2012, 2018; Revaz et al. 2016), a fully parallel chemo-dynamical Tree/Smoothed Particle Hydrodynamics(SPH) code based on GADGET-2 (Springel 2005). In addition to the collision-less dark matter, GEAR treats the baryonic physics by including gas cooling, redshift-dependent UV-background heating, star formation, and stellar feedback. We briefly summarize its essential elements below.

Radiative cooling. Gas cooling is computed using the GRACKLE library (Smith et al. 2017). In addition to primordial gas cooling, it includes metal-line cooling scaled according to the gas metallicity. GRACKLE also includes UV-background radiation based on the prediction from Haardt & Madau (2012). Hydrogen self-shielding against the ionizing radiation is incorporated by suppressing the UV-background heating for gas densities above $n_H = 0.007\,\mathrm{cm^{-3}}$ (Aubert & Teyssier 2010). It should be noted that while GRACKLE includes H_2 and dust, we did not consider those spices in our model. At the current resolution of our simulations, the maximal density reached by the gas is about 100 atom $\mathrm{cm^{-3}}$, below the regime where dust starts to play a major role as a coolant (see e.g., Omukai et al. 2005). Moreover, with such resolution, at a density of $n_H = 1$ atom $\mathrm{cm^{-3}}$ the Jeans length is not resolved for gas with a temperature below 100 K. In this regime, as described below, we relied on the pressure-floor method, and thus did not solve the very dense and cold gas in a fully self-consistent way. The lack of H_2 cooling could slightly delay the formation of the first stars. The impact on our results was nevertheless very mitigated by the very large variety of star formation histories that our models sampled. We leave further improvements to our models, containing the treatment of dust and its impact on the cooling process (e.g., as done in Bekki 2015; McKinnon et al. 2017; Aoyama et al. 2019; Chiaki et al. 2020) for future studies. Also, it is worth mentioning that while the cosmic UV-background radiation was included, our physical model did not include ionizing photons emitted from the in situ massive stars.

---

2 One resolution level l corresponds to $N = (2^l)^3$ particles in the full cosmological box. The particle mass is thus decreased by a factor of two between two levels.

Star formation. Star formation is modeled using the stochastic prescription proposed by Katz (1992) and Katz et al. (1996) that reproduces the Schmidt law (Schmidt 1959). This classical recipe is supplemented by a modified version of the Jeans pressure floor through adding a nonthermal term in the equation of state of the gas. The purpose of this modification is to avoid any spurious gas fragmentation (Truelove et al. 1997; Bate & Burkert 1997; Owen & Villumsen 1997). In dense regions where the system is dominated by unresolved physics, we adopted a star formation density threshold based on the Jeans polytrope, directly correlating the SPH smoothing length with density:

$$\rho_{\mathrm{SFR},i} = \frac{\pi}{4} G^{-1} N_{\mathrm{Jeans}}{}^{-2/3} h_i^{-2} \left(\gamma \frac{K_{\mathrm{B}}}{\mu m_{\mathrm{H}}} T + \sigma_i{}^2 \right). \tag{1}$$

Here G is the universal gravitational constant and γ the adiabatic index of the gas fixed to 5/3. h_i and ρ_i are the SPH smoothing length and density, respectively. σ_i is the velocity dispersion of the gas particle i, summed over the neighboring particles in the Jeans mass. N_{Jeans} is the ratio between the Jeans mass and the SPH mass resolution of particle i, and is fixed to ten. T corresponds to the gas temperature. In this work, we supplemented this temperature-dependent threshold with a constant density threshold $\rho_{\mathrm{SFR},c}$, which prevents stars from forming in the cold low-density gas regions. The importance of this threshold and its value is discussed in Sect. 4.1. Above $\max(\rho_{\mathrm{SFR},c}, \rho_{\mathrm{SFR},i})$, stars may form with a star formation efficiency of 1%.

Depending on the metallicity of the gas particle from which it will form, the stellar particle is composed of either Pop II or Pop III stars. Pop III stars only form from gas particles with [Fe/H] lower than a critical value, [Fe/H]$_c$. Above this gas metallicity, only Pop II stars will form. For our fiducial model, we set [Fe/H]$_c$ equal to -5, as above this metallicity fine-structure line cooling of atomic carbon and oxygen (Bromm & Loeb 2003), but also dust grain emission (Omukai et al. 2005), lead to gas fragmentation and the formation of low-mass stars. As presented below, we assumed Pop III stars to be massive stars. Thus, in our model, all stars more metal poor than [Fe/H]$_c$ explode in a few Myr, and will be absent at redshift $z = 0$. In Sect. 4.6.1, we discuss the impact of modifying [Fe/H]$_c$. Because [Fe/H]$_c$ is not reached in all gas particles at the same time, these Pop II and Pop III stars can potentially coexist at a given time step.

3.2. Initial mass function

Pop III stars are predicted to have specific IMF and yields (e.g., Iwamoto 2005; Heger & Woosley 2010; Bromm 2013), and we thus updated GEAR with the capability to fully distinguish the properties of Pop II and Pop III stars. In our approach, each stellar particle represents a single stellar population (SSP) that is associated with its own IMF and stellar yields.

For both Pop II and Pop III stars, the IMF is described by the probability function $\Phi(m)$, where $\Phi(m)\,\mathrm{d}m$ gives the fraction of stars born in the mass range $[m, m + \mathrm{d}m]$. In our implementation, $\Phi(m)$ is defined by a set of power laws with the slope $\alpha(m)$ depending on the mass interval :

$$\Phi(m) = \frac{m^{\alpha(m)}}{\beta}, \tag{2}$$

with β the normalization constant that ensures the integral of $\Phi(m)$ over the entire mass range considered to be one.

Pop II. We used the revised IMF of Kroupa (2001), including systematic bias due to unresolved binaries in observations, previously implemented in GEAR with

$$\alpha(m) = \begin{cases} 0.7, & \text{if} \quad m \in [0.05\,M_\odot, 0.08\,M_\odot] \\ -0.8, & \text{if} \quad m \in [0.08\,M_\odot, 0.50\,M_\odot] \\ -1.7, & \text{if} \quad m \in [0.50\,M_\odot, 1.00\,M_\odot] \\ -1.3, & \text{if} \quad m \in [1.00\,M_\odot, 50.0\,M_\odot] \end{cases} \tag{3}$$

The minimal and maximal stellar masses were set to 0.05 and $50\,M_\odot$, respectively.

Pop III. Adopting a different IMF for Pop III than that of Pop II stars was motivated by the fact that no metal-free star has been observed so far, an indication that there are no, or very few, long-lived primordial low-mass stars. It is thus expected that the majority of Pop III stars are massive enough to explode as CCSNe and disappear (Larson 1998; Bromm et al. 1999; Karlsson et al. 2008). For Pop III stars, we adopted a simple power-law IMF with a unique slope $\alpha = -1.3$, similar to the one of the most massive Pop II stars. The minimal mass $M_{\min}$ was set to $13\,M_\odot$. It is worth mentioning that $M_{\min}$ was above the lower supernova mass limit. The maximal mass $M_{\max}$, when PISNe were not considered, was set to $140\,M_\odot$. When PISNe were included, $M_{\max}$ was set to $300\,M_\odot$.

3.3. Stellar feedback: Yields and energy

The stellar lifetimes were mass- and metallicity-dependent, following Kodama & Arimoto (1997, priv. comm.). At each time step, the exploding supernovae were randomly selected following the random discrete IMF sampling scheme of Revaz et al. (2016), in which the IMF was considered as a probability distribution. We calculated the number of potentially exploding supernovae $N_{\mathrm{SN},i}(t, \Delta t)$ at each time step. For each stellar particle i, this number depended on the lifetimes of stars it contained. The integer part of $N_{\mathrm{SN},i}(t, \Delta t)$ was assumed to explode as supernovae. The remaining fractional value was compared to a random variable χ between zero and one. If the fractional part was above χ, the number of exploding supernovae was increased by one, otherwise, the number of dying stars was equal to the integer part.

As to the massive stars that lead to supernovae explosion, our previous works (Revaz & Jablonka 2012, 2018; Nichols et al. 2014, 2015; Revaz et al. 2016; Harvey et al. 2018; Hausammann et al. 2019; Sanati et al. 2020), only included CCSNe, with the yields of Tsujimoto et al. (1995), for stellar masses up to $70\,M_\odot$. We had not yet considered zero metallicity stars. Therefore, for the purpose of the present study, for which we do need Pop III stars, we introduced the metallicity-dependent yields of Nomoto et al. (2013, hereafter N13), from [M/H] $= -\infty$ to 0.4[3]. Figure 2 summarizes the yields and feedback energies of Pop II and Pop III stars, as considered in this study and further described below.

Pop II. As indicated above, we used the yield tables of N13, taking into account the analysis and conclusions of Kobayashi et al. (2020). The N13 yields were provided at [M/H] $= -\infty, -1.3, -0.7, -0.4, 0, 0.4$. As GEAR only treats metallicity-independent yields, we interpolated the N13 tables at [M/H] $= -2.5$, which globally corresponds to the average stellar metallicity of UFDs. The result of this interpolation is seen with thick blue lines in Fig. 2. Below $40\,M_\odot$ the values of the N13 table for the two lowest metallicities, [M/H] $= -\infty$ and

---

[3] $\quad \mathrm{[M/H]} = \log_{10}\left(\frac{M_{\mathrm{metals}}}{M_{\mathrm{H}}}\right)_\star - \log_{10}\left(\frac{M_{\mathrm{metals}}}{M_{\mathrm{H}}}\right)_\odot.$
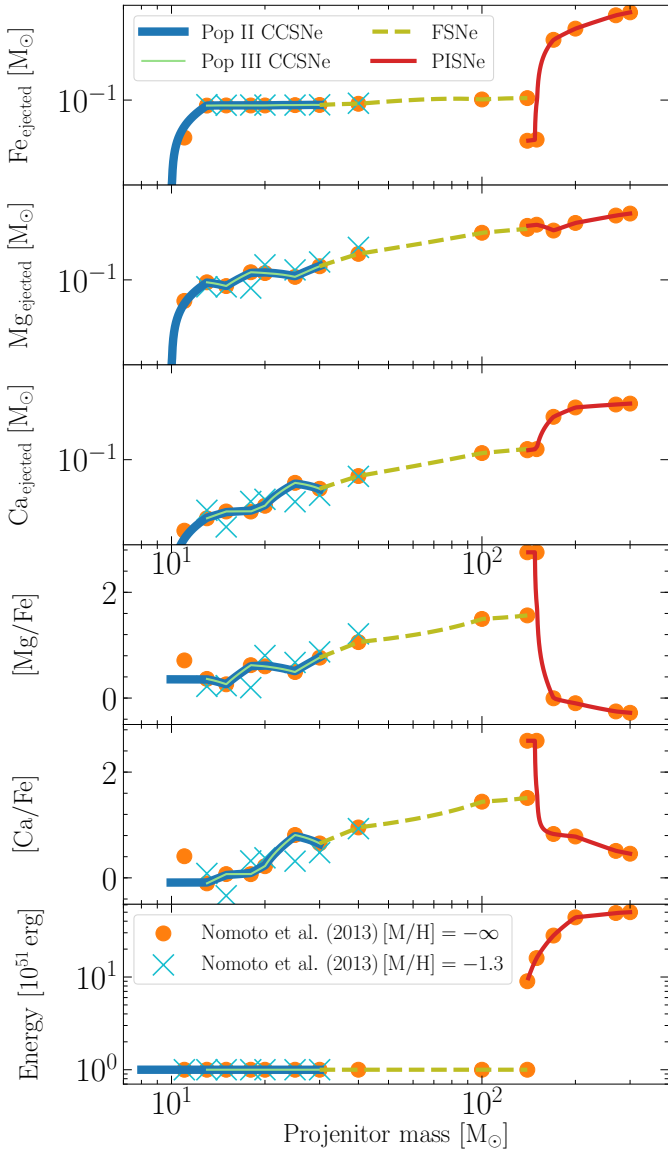
Fig. 2. Comparison of the yields of iron, magnesium, and calcium as a function of stellar masses from N13. The corresponding abundance ratios [Mg/Fe] and [Ca/Fe], and the supernovae explosion energies are also shown. The N13 table values for stars with $M_{\mathrm{proj}} < 30\,M_\odot$ at $[\mathrm{M/H}] = -\infty$ and -1.3 are shown with orange circles and blue crosses, respectively. The thick blue and thin green curves respectively show the adopted, interpolated values used in this work for the Pop II and Pop III stars. The dashed yellow lines indicate the mass ranges of stars ending as FSNe. The red curves show the yields and energies for PISNe with $M_{\mathrm{proj}} > 140\,M_\odot$ examined in this work from N13.

$[\mathrm{M/H}] = -1.3$, are shown with orange circles and blue crosses, respectively. The proximity of these two sets of values (barely visible for the iron ejecta) supports our interpolation.

As in Kobayashi et al. (2020), only supernovae progenitors less massive than $30\,M_\odot$ contribute to the chemical enrichment. Above this mass, models predict that the formation of supernova progenitors is quite unlikely (Smartt 2009). We thus ignored the explosion of the stars that are referred to as failed supernovae (FSNe) in the text. For all Pop II CCSNe with mass less than $30\,M_\odot$ the energy released per explosion was fixed to the fiducial value of 10^{51} erg.

SNeIa. Stars with initial masses between 3 and $8\,M_\odot$ can be progenitors of SNeIa, following the model of binary systems of Kobayashi et al. (2000). The SNeIa yields were taken from Tsujimoto et al. (1995). Their explosion energy was fixed to 10^{51} erg.

Pop III. The yields and energies of Pop III stars are those of the metal-free ($[\mathrm{M/H}] = -\infty$) stars provided by N13. In the mass range 13–$30\,M_\odot$, shown with thin green lines in Fig. 2, Pop III stars explode as CCSNe with feedback energy 10^{51} erg. As the uncertainties in the physics of black hole formation from the first massive stars obscure the upper limit of CCSNe, for Pop III with initial masses between 30 and $140\,M_\odot$ (dashed yellow lines), we examined two different final fates: (i) normal CCSNe with nucleosynthesis yields from N13, and (ii) FSNe that, according to Kobayashi et al. (2020), end as black holes. In this latter case, no energy nor yields are ejected.

In the PISNe regime for stars above $140\,M_\odot$, yields were also taken from the N13 table, corresponding to the model predictions of Umeda & Nomoto (2002). In this mass regime, shown with red lines, supernova energy varies from 9 to 50×10^{51} erg for the most massive stars with $300\,M_\odot$.

3.4. Stellar feedback

Exploding supernovae release both synthesized chemical elements and energy into the ISM. In particle-based codes, this is modeled by distributing these quantities on the neighboring particles, whose definition and identification are not trivial. We describe here the two methods implemented in GEAR and discuss their advantages and drawbacks. The impact of these schemes, as well as their parameters, are studied in Sect. 4.1.

Injecting into the SPH kernel. The standard method consists in releasing the ejecta into the effective number of neighboring particles, N_{eff}, which is determined by the implemented SPH scheme. In the modern conservative formulation of SPH (Springel & Hernquist 2002; Price 2012; Springel 2010; Hopkins 2013) as used by GEAR, a differenciable relation between the volume of a particle (or equivalently its density) and its smoothing length is required. The most natural one is to consider that mass in the kernel volume of a particle remains constant, equal to $\bar{N}_{\mathrm{ngb}}$ times the mass m_i of particle i. The parameter $\bar{N}_{\mathrm{ngb}}$ (fixed to 50) sets the average number of neighboring particles defined by:

$$\frac{4}{3}\pi h_i^3 \rho_i = m_i \bar{N}_{\mathrm{ngb}}, \tag{4}$$

where the smoothing length h_i and density ρ_i of particle i are related to the volume V_i by $\rho_i = m_i/V_i$. Defining V_i through the summation over neighboring particles, via the kernel function W[4]:

$$V_i = \frac{1}{\sum_{j=1}^{N_{\mathrm{eff}}} W(|\boldsymbol{r}_i - \boldsymbol{r}_j|, h_i)}, \tag{5}$$

one can explicitly express $\bar{N}_{\mathrm{ngb}}$:

$$\bar{N}_{\mathrm{ngb}} = \frac{4}{3}\pi h_i^3 \sum_{j=1}^{N_{\mathrm{eff}}} W(|\boldsymbol{r}_i - \boldsymbol{r}_j|, h_i). \tag{6}$$

This latter relation shows that the effective number of neighboring particles N_{eff} is different from the fixed parameter

---

[4] It should be noted that there are different possibilities to express V_i. See Hopkins (2013).

Table 1. Parameters of the simulations.

(1) Model description Model ID	(2) Mass range [$M_\odot$] Pop II Pop III	(3) IMF slope(α) Pop II Pop III	(4) CCSNe Pop II Pop III	(5) FSNe Pop II Pop III	(6) PISNe M [$M_\odot$] E [10^{51} erg]	(7) Critical metallicity [Fe/H]$_c$	(8) Density threshold $\rho_{SFR,c}$ [atom/cm^3]	(9) Metal ejection radius R_{ejecta}[pc]
Models with only Pop II stars								
PopII	[0.05, 50] −	[0.7, −0.8, −1.7, −1.3] −	[8, 30] −	[30, 50] −	– −	–	5	300
Models with Pop III stars								
PopII+PopIII(CCSNe)	[0.05, 50] [13, 140]	[0.7, −0.8, −1.7, −1.3] −1.3	[13, 30] [13, 140]	[30, 50] −	– −	−5	5	300
Models with Failed SNe								
PopII+PopIII(CCSNe+FSNe)	[0.05, 50] [13, 140]	[0.7, −0.8, −1.7, −1.3] −1.3	[13, 30] [13, 30]	[30, 50] [30, 140]	– −	−5	1 , 5	h_i , 300
Models with PISNe								
PopII+PopIII(CCSNe+FSNe)+PISNe	[0.05, 50] [13, 300]	[0.7, −0.8, −1.7, −1.3] −1.3	[13, 30] [13, 30]	[30, 50] [30, 140]	[140, 300] [9, 50]	−5	5	300
Models with decreasing PISNe feedback energy								
PopII+PopIII(CCSNe+FSNe)+PISNe(E$_{51}$)	[0.05, 50] [13, 300]	[0.7, −0.8, −1.7, −1.3] −1.3	[13, 30] [13, 30]	[30, 50] [30, 140]	[140, 300] 1	−5	5	300
Models with increasing critical metallicty								
PopII+PopIII(CCSNe+FSNe)+PISNe(Z-4)	[0.05, 50] [13, 300]	[0.7, −0.8, −1.7, −1.3] −1.3	[13, 30] [13, 30]	[30, 50] [30, 140]	[140, 300] [9, 50]	−4	5	300

Notes. Columns are as follows: 1) Model description and ID. 2) Mass range of Pop II and Pop III stars. 3) IMF slopes. 4) Mass range of Pop II and Pop III stars that fall into a CCSNe, for which a feedback energy of 10^{51} erg is assumed. 5) Mass range of Pop II and Pop III stars that fail to explode and fall into black holes. 6) Mass range of Pop III stars forming PISNe and the feedback energy released after their explosion. 7) Critical gas metallicity below which only Pop III stars form. Above this metallicity only Pop II stars form. 8) Density threshold in gas clouds, $\rho_{SFR,c}$, required for the onset of star formation. 9) Radius, R_{ejecta}, around each supernova in which metals are ejected.

$\bar{N}_{ngb}$, and that N_{eff} strongly depends on the distribution of the neighboring particles through $W(|\mathbf{r}_i - \mathbf{r}_j|, h_i)$. We thus expect this number to be different for each time ejection event.

A particular case worth mentioning is when two supernovae explode subsequently in the same region. In this case, the second supernova explodes in the Sedov cavity created by the first explosion. Assuming an initial homogeneous ISM, for the first explosion, the SPH scheme will find $N_{eff} \cong \bar{N}_{ngb}$, while for the second supernova, N_{eff} will largely exceed $\bar{N}_{ngb}$. Indeed, as neighboring particles can only be found at larger distances, each one of them will have a low $W(|\mathbf{r}_i - \mathbf{r}_j|, h_i)$. In this case, a large number of particles have to be considered in the sum of Eq. (6), in order to guarantee the convergence toward $\bar{N}_{ngb}$. Consequently, this scheme can potentially cause the ejection of metals into unrealistic and uncontrolled distances. Moreover, it is strongly resolution dependent as h_i scales with the spatial resolution. We further discuss the consequences of using this scheme in Sect. 4.1.

Injecting into a constant volume. An alternative to the previous scheme is to distribute the ejecta among particles that reside in a constant volume. This volume is defined by a sphere of radius R_{ejecta} around the exploding star. In addition to avoiding the bias of ejecting into the SPH volume, the method is motivated by reproducing the effective extent of a supernova explosion. The adiabatic supernova explosion in a homogeneous medium must follow the Sedov-Taylor solution (Sedov 1959; Taylor 1950). The mean pressure $\bar{P}$ within the blast wave falls off rapidly with radius. This means that the expansion of the supernova shock must follow a power law that, at the end of the blast wave phase, using typical values for supernova explosions, is approximately equal to:

$$R_{ejecta} \approx 300\,\text{pc} \left(\frac{E}{10^{51}\,\text{erg}} \right)^{1/3} \left(\frac{\bar{P}}{4\times10^{-13}\,\text{dyne cm}^{-2}} \right)^{-1/3}, \quad (7)$$

where we assume that an amount of energy E is instantaneously ejected into the ambient medium of uniform density ρ, characterized by the adiabatic index $\gamma = \frac{5}{3}$ (Ostriker & McKee 1988). We thus supplemented GEAR with the possibility to eject metals in a constant physical volume (not a comoving volume) using R_{ejecta} that we chose to be 300 pc.

Delayed cooling and mixing. In both schemes mentioned above, deposited energy and metals are weighted by the kernel $W(|\mathbf{r}_i - \mathbf{r}_j|, h_i)$. To avoid an instantaneous radiation of the injected energy, we used the delayed cooling method, which consists in disabling gas cooling for a short period of time (Stinson et al. 2006), here taken as 5 Myr. Only a fraction ϵ of the supernova energy is deposited into the ISM. We assumed $\epsilon = 0.2$, and hence ~20% of the released energy impacted the ISM effectively. The released chemical elements were further mixed in the ISM using the Smooth metallicity scheme (Okamoto et al. 2005; Tornatore et al. 2007; Wiersma et al. 2009). In short, the chemical abundance of a gas particle results from the convolution of the metal content of its neighbors.

3.5. Models

We explored the impact of the different parameters listed in Table 1. The stellar range of properties, as well as the model ID are provided in the first column.

Models including only Pop II stars are our references, hereafter denoted as PopII. The mass range of Pop II and Pop III stars is given in the second column. As described in Sect. 3.2, the IMF slope of the Pop II stars changes with the stellar mass range (Kroupa 2001). For the Pop III stars, we adopted a fixed slope of −1.3, similar to the most massive of the Pop II stars.

The mass range of CCSNe is shown in the forth column. In PopIII(CCSNe), all Pop III stars explode as a normal CCSNe, while in PopIII(CCSNe+FSNe), stars in mass range [30−140] $M_\odot$ fail to explode and end as black holes. The mass

range of these FSNe (see Sect. 3.3), ejecting no energy nor yields, is given in the fifth column.

For massive Pop III stars ([140−300] $M_\odot$) exploding as PISNe, we explored the two feedback energy prescriptions indicated in the sixth column: (i) in our reference PISNe model, the highly energetic feedback increases from 9 to 50×10^{51} erg, according to the initial mass of the star (see N13, for more details); (ii) in the PISNe(E_{51}) model, we explored a less energetic feedback. In this model, the supernova energy is independent of the progenitor mass and equal to the energy ejected by the normal CCSNe (i.e., 10^{51} erg).

The critical gas metallicity [Fe/H]$_c$, below which only the first massive and metal-free stars may form, is shown in the seventh column. When the metallicity of gas is higher than this critical value, low-mass long-lived Pop II stars can form. In model PISNe(Z-4), we increased [Fe/H]$_c$ from −5, of our reference PISNe model, to −4.

In model PopIII(CCSNe+FSNe), we explored the impact of the gas density threshold, $\rho_{\mathrm{SFR},c}$, required for the onset star formation. In this model we also explored the radius R_{ejecta} around each supernova in which metals were ejected (see Sect. 3.4). These two parameters are given in columns eight and nine. In all other models, these parameters were fixed to 5 atom/cm^3 and 300 pc, respectively.

3.6. Extraction of the observables, luminosity, and metallicity

Both the galaxy luminosity and metallicity were extracted from the low-mass long-lived stars traced by the Pop II stellar particles. All Pop III stars explode, and therefore do not contribute to the final stellar mass budget, luminosity, or metallicity of their host galaxy. The line-of-sight (LOS) stellar velocity dispersion, σ_{LOS}, was calculated for seven different lines of sight inside a 1 kpc cylindrical radius. The value quoted for each galaxy represents the mean of these values.

The galaxy V-band luminosity (L_V) was obtained by computing the luminosity of individual Pop II stellar particles located within the 90% of the halo's light radius. Their mass was converted into luminosity using the stellar population synthesis model of Vazdekis et al. (1996), computed with the revised Kroupa (2001) IMF. Where necessary, the luminosities were inter- and extra-polated in age and metallicity using a bivariate spline. We ignored dust absorption, as the stellar metallicity of galaxies studied here were [Fe/H] < −2. Indeed, at these low metallicities, the dust-to-gas ratio was lower than 10^{-5} (see e.g., Fig. 1 in Hirashita 2015).

Because of the limited number of stars formed in UFDs, the choice of a representative "mean" [Fe/H] is not a trivial task. As the metallicity distribution is sparsely sampled computing the mode, the peak of the metallicity distribution function can lead to large uncertainties. In Appendix A, we propose a method to determine the mode based on a fitting analytical formula derived from a simple chemical evolution model. The error on [Fe/H] was taken at the maximum of the errors obtained by the different methods, peak and mode. All chemical abundances were calculated with respect to the solar abundances of Anders & Grevesse (1989).

4. Results

Table 2 summarizes the main properties of the simulated halos at $z = 0$ for our fiducial model, which includes only Pop II stars. The halo IDs are those of Revaz & Jablonka (2018, see their Table 1), while the simulations were run with the physics described in this paper. Table 2 provides the galaxy total V-band luminosity L_V, total stellar mass $M_\star$, virial mass M_{200}, virial radius R_{200}, mean stellar velocity dispersion σ_{LOS}, and the peak value of the stellar metallicity distribution function [Fe/H]. The halo h177 hosts a Sextans-like classical dwarf galaxy (see next section). The other 18 halos host dwarf galaxies with luminosities equal to or lower than $10^5 L_\odot$ and are considered to be UFDs.

4.1. Model calibration and validation

While in Revaz & Jablonka (2018) we could successfully reproduce the observed properties of the Local Group classical dwarf galaxies, in the present study, not only did we use different nucleosynthesis tables, but we also increased the effective resolution of the simulations by one order of magnitude. Therefore, before taking up the challenge of UFDs, we first needed to calibrate our physical reference model. To this end, we focused on the halo h177, representative of a Sextans-like galaxy. Indeed Sextans is one the least massive classical dwarf galaxies and it benefits from extensive medium- and high-resolution spectroscopy, including abundance ratios over a wide metallicity range. This enables a proper and detailed comparison with the models.

Compared to Revaz & Jablonka (2018), the supernova efficiency was increased by a factor of two, from $\epsilon = 0.1$ to 0.2. With this setting, our Sextans-like model h177 had a total V-band luminosity of about $3 \times 10^5 L_\odot$, only 25% lower than the observed value (McConnachie 2012). Star formation fully ended after 3 Gyr, in agreement with the star formation history derived from color magnitude diagrams (Lee et al. 2003, 2009; Bettinelli et al. 2018). The peak metallicity (−1.97) and the velocity dispersion (∼6 km s^{-1}) were both in agreement with the values derived from the calcium triplet (Battaglia et al. 2011).

In a ΛCDM cosmological framework, dwarf galaxies, and particularly UFDs, are dominated by metal-poor stars, which form in clumps in the initial mini-halos. After a few hundred Myrs, these clumps merge and form the final galaxy. This sequence is shown in Fig. 3, which displays the build-up history of h177 from 190 to 430 Myr, a period that corresponds to the first intense phase of the star formation in this system. For illustration purposes, we show here the model PopII+PopIII(CCSNe+FSNe), which includes Pop III stars, but the sequence is identical in all models. The surface density of the gas is color-coded in blue. The Pop III and Pop II stars are respectively shown by yellow and red symbols. The hierarchical assembly of the dwarf is clearly seen; UFDs follow the same hierarchical pattern.

From Fig. 3, it is easy to understand that, besides the supernovae feedback energy, two other parameters are crucial to the evolution of the model dwarfs: the density threshold, $\rho_{\mathrm{SFR},c}$, above which stars form, and the size of the regions over which the supernovae ejecta are distributed. These two parameters are constrained by considering the trend and dispersion of the stellar abundance ratios, considering the α-elements Mg and Ca, as a function of metallicity ([Fe/H]).

The star formation density threshold. In our previous studies, star formation was enabled when the gas density of particle i was larger than $\rho_{\mathrm{SFR,i}}$ (defined by Eq. (1)). Indeed, while the probability is low, in principle, it is possible that stars form from low-density gas (i.e., lower than 1 atom cm^{-3}). Together with the gas surface density, Fig. 3 traces the metallicity of the gas ([Fe/H]$_{\mathrm{gas}}$) in orange. This allows us to follow the regions that are enriched by the supernovae ejecta.

Table 2. Global properties of the 19 halos simulated when only Pop II stars are considered.

Halo ID	L_V	$M_\star$	M_{200}	R_{200}	$r_{1/2}$	$\sigma_{\rm LOS}$	[Fe/H]
	$[10^5 \, L_\odot]$	$[10^5 \, L_\odot]$	$[10^9 \, M_\odot]$	[kpc]	[kpc]	[km s^{-1}]	[dex]
h177	3.07	10.28	0.51	25.9	0.26	5.7	−1.97
h063	1.41	3.73	2.41	41.8	0.51	6.4	−2.31
h190	1.31	3.62	0.47	24.2	0.28	4.8	−2.15
h187	1.04	2.86	0.43	23.4	0.25	5.3	−2.27
h166	0.71	1.92	0.64	26.9	0.22	5.0	−2.36
h112	0.63	1.68	1.18	32.9	0.43	6.9	−2.79
h152	0.53	1.42	0.80	28.9	0.38	6.4	−2.82
h113	0.41	1.08	1.14	32.5	0.53	7.4	−3.07
h184	0.30	0.81	0.61	26.4	0.28	4.8	−2.69
h245	0.17	0.45	0.39	22.7	0.32	6.0	−3.24
h259	0.20	0.53	0.39	22.7	0.21	5.1	−2.74
h226	0.15	0.38	0.44	23.6	0.25	4.8	−2.87
h277	0.10	0.26	0.36	22.2	0.42	5.3	−3.03
h249	0.07	0.18	0.39	22.8	0.28	5.7	−3.18
h315	0.05	0.14	0.33	21.6	0.51	5.6	−3.16
h323	0.05	0.13	0.27	20.1	0.21	4.0	−3.07
h170	0.05	0.13	0.62	26.5	0.41	5.0	−3.56
h273	0.04	0.09	0.39	22.7	0.32	4.1	−3.53
h291	0.02	0.04	0.33	21.4	0.18	5.5	−3.69

Notes. The first column gives the halo ID following Revaz & Jablonka (2018). L_V is the V-band luminosity and $M_\star$ the stellar mass. M_{200} is the virial mass, i.e., the mass inside the virial radius R_{200}. $r_{1/2}$ is the half-light radius, $\sigma_{\rm LOS}$ is the line-of-sight velocity dispersion, and [Fe/H] is the peak value of the metallicity distribution function.
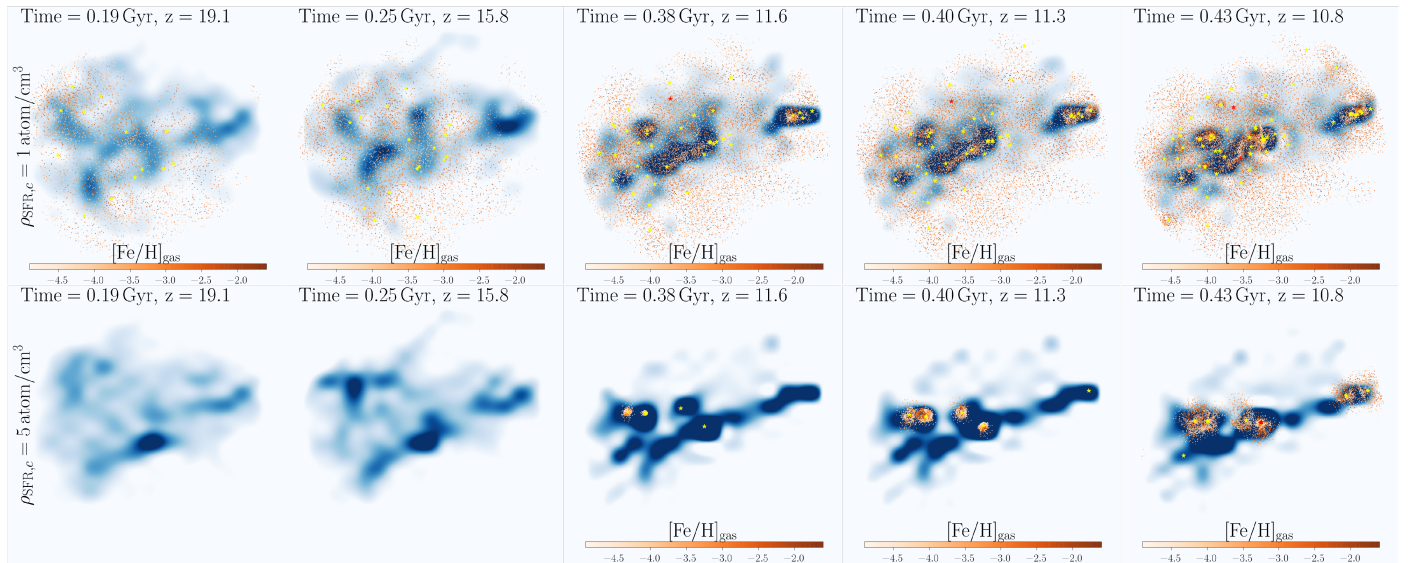
Fig. 3. Metal enrichment of the Sextans-like model galaxy, h177, from 190 to 430 Myr. The surface density of gas is color-coded in blue. After each supernova explosion, shown in yellow and red for Pop III and Pop II stars, respectively, yields were ejected, using the SPH kernel scheme. The metallicity of gas, color-coded in orange, was determined by the amount of Fe received from the supernovae ejecta. The star formation density threshold $\rho_{\rm SFR,c}$ was set to 1 and 5 atom cm^{-3} in the top and bottom panels, respectively.

In the top panel of Fig. 3, $\rho_{\rm SFR,c}$ is set to 1 atom cm^{-3} (1.67×10^{-24} g cm^{-3}). Star formation starts at ∼100 Myr in mini-halos of low gas density. Because their reservoir of gas is small, only one generation of stars can occur in each of the mini-halos. Metals ejected by Pop III stars are dispersed over large regions, of up to ∼500 pc h^{-1} physical radii. Consequently, only a small fraction of these metals are locked into the second generation of stars.

Not only this, but random sampling of IMF in each mini-halo naturally leads to a various amount of metals released in differ-

ent mini-halos. The fact that mixing is very inefficient between them results in a large scatter in the abundance ratios ([Mg/Fe], [Ca/Fe]) at a fixed metallicity, above the dispersion measured in the observations.

In the bottom panel of Fig. 3, $\rho_{\rm SFR,c}$ is set to 5 atom cm^{-3} (8.36×10^{-24} gr cm^{-3}), that is to say five times the value used in the simulation shown in the top panel. One immediate impact is the delayed onset of star formation by about ∼200 Myr, as gas particles need time to merge and sufficiently increase the gas density.

Another consequence of the increase in $\rho_{\text{SFR},c}$ is that, being more massive, the gas reservoirs enable the formation of more than one generation of stars in each of the mini-halos. The ejecta of the Pop III stars increase the metallicity of the gas to about [Fe/H] = −3, then multiple Pop II stars form from this enriched gas in the same mini-halo, inheriting the chemical imprints of the Pop III stars. Because star formation is more spatially concentrated and the number of gas particles higher, the dilution of the metals in the ISM is reduced. Thus, in order to avoid the dispersion of metals in a large unrealistic volume, such as occurs with $\rho_{\text{SFR},c} = 1 \, \text{atom cm}^{-3}$, we set the star formation density threshold to 5 atom cm^{-3}. This choice resulted from our investigation of a range of values, from 2 to 10 atom cm^{-3}, and best reproduced the observed scatter in [Mg/Fe] and [Ca/Fe].

Spatial extent of the ejection of metals. As mentioned above, the extent of the region polluted by the supernova ejecta impacts the chemical abundance pattern of the gas particles, later imprinted in the stars. This includes metallicity and abundance ratios. We investigated two ejection schemes with the aim of reproducing the observed galaxy stellar metallicity distributions, as well as the trends and scatter of the α-element abundance ratios with metallicity.

We performed a statistical analysis of the properties of the regions reached by the supernovae explosions. The size of the polluted region R_{ejecta} is defined as the distance to the furthest particle receiving the supernova ejecta. Figure 4 presents the 2D histogram of R_{ejecta} and time, during the first 1.2 Gyr of evolution of the halo h177. Each point of the diagram corresponds to the number of supernova explosions, which at a given time enriched the ISM up to R_{ejecta}. Blue corresponds to the case where the metals were ejected into the SPH kernel, while red stands for the case where the ejecta were distributed in a volume of constant *maximum* radius of 300 pc (see Sect. 3.4 for the details of both methods). Before applying any additional constraints to the metal ejection schemes, the SPH radius was generally smaller than 100 pc, with the consequence that a large part of the gas particles were left in their pristine state, slowing down the metal enrichment of the ISM. Therefore, from the point of view of the galaxy mean metallicity, increasing the diffusion radius of the metals is more efficient. As to the abundance ratios, the general trend of [α/Fe] versus [Fe/H] was successfully reproduced by both implementations, in particular the position of the so-called knee, which indicates the time when the SNeI ejecta significantly contribute to the metal content of the ISM.

Generally, the size of the polluted regions are much smaller in the SPH kernel ejection scheme than in the constant volume method. In dense gas clouds, where the $\bar{N}_{\text{ngb}}$ neighboring particles are found at small distances, R_{ejecta} may be as small as 10 pc h^{-1}. As expected, in the constant volume ejection scheme, the size of the maximum polluted region is 300 pc on average. This radius can occasionally be smaller in cases of very isolated gas clouds.

The small ejecta radii induced by the SPH scheme leads to less efficient metal mixing and consequently larger scatter in the abundance ratios of α-elements, as seen in Fig. 5, which displays the stellar [Mg/Fe] and [Ca/Fe] as a function of [Fe/H]. The two settings for metal ejection, the SPH kernel and the constant volume (see Table 1), are shown in blue and green, respectively. For the sake of a fair comparison with the observations, we added random uncertainties to the model abundances, following a normal distribution with a standard deviation of 0.1 dex. Using the same yields, Kobayashi et al. (2020) semi-analytical models produce a [Mg/Fe] plateau 0.2 dex higher than the obser-
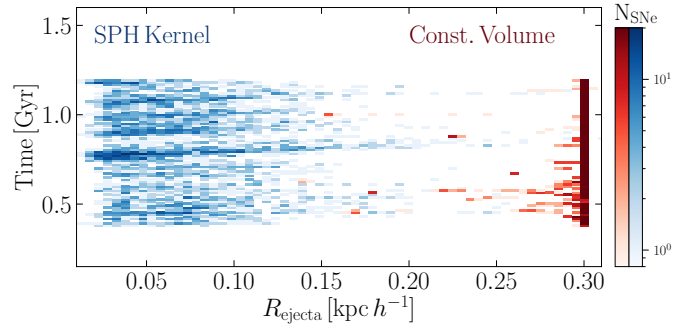

Fig. 4. Comparison between the two different methods for metal ejection implemented in GEAR. Injection into the SPH kernel and into a constant volume, as described in Sect. 3.4, are shown in blue and red, respectively, for the Sextans-like model h177. Both models are plotted as the 2D histogram of the extent of the polluted region R_{ejecta}, and the time at which the explosion occurs, limited to the first ~1 Gyr of the simulation. Pixels are color-coded by the number of supernova events. In each event, R_{ejecta} is defined as the distance to the furthest particle that receives the supernova remnants.
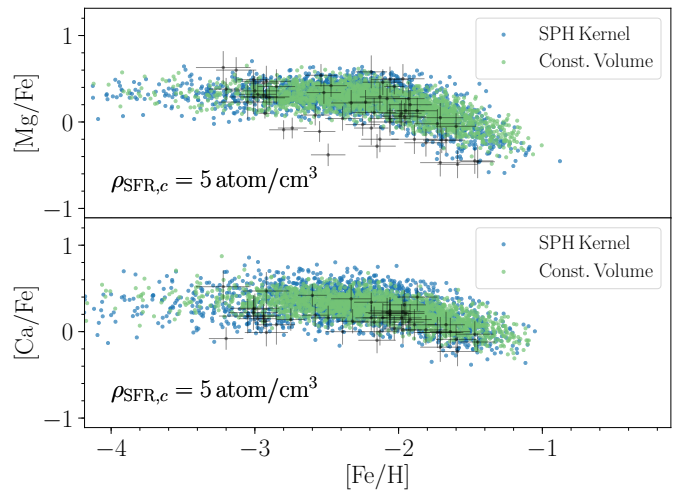

Fig. 5. [Mg/Fe] and [Ca/Fe] versus [Fe/H] for stars formed in the two different methods for metal ejection, compared to the observed stars. Metal ejection into the SPH kernel and into a constant volume are shown in blue and green, respectively. Observational values for Sextans stars are shown with black dots and error bars. The data were taken from Shetrone et al. (2001), Kirby et al. (2010), Tafelmeyer et al. (2010), Theler et al. (2020), and Lucchesi et al. (2020).

vations. This small difference is interpreted as uncertainties in the stellar yields. We therefore shifted our model [Mg/Fe] by −0.2 dex in all subsequent figures involving abundance ratios. Figure 5 shows that the scatter induced by the SPH scheme leads to a larger scatter than observed[5]. We therefore adopted a constant R_{ejecta} for our simulations as it leads to an improved chemical mixing.

4.2. New yields for Pop II stars

Compared to our previous work, the resolution, the yields, and the injection scheme changed. In a first step, we simulated our

---

[5] It should be noted that the dispersion in the observed [Mg/Fe] is larger than in the observed [Ca/Fe] simply because the number of lines used for the derivation of the magnesium was much smaller than for the calcium.
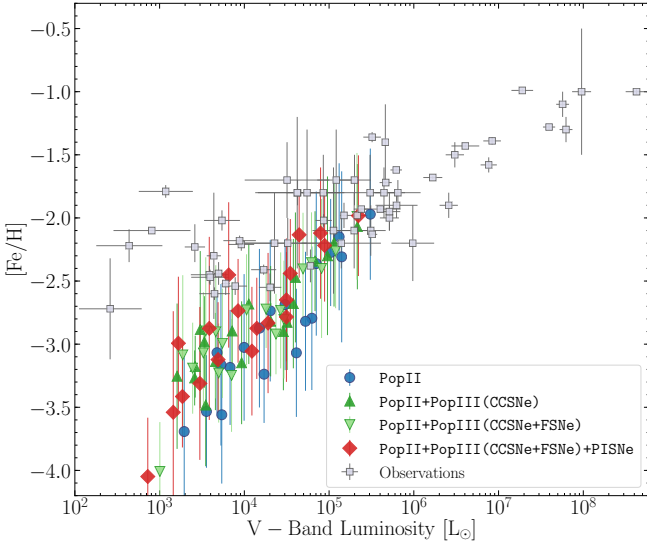
Fig. 6. Comparison of the luminosity-metallicity relation for dwarfs and UFDs, between Local Group observations and simulations. The gray squares represent our Local Group sample (see text for details). Colored points stand for different simulations described in Table 1: The reference PopII model is in blue points, model PopIII(CCSNe) including yields from both massive and low-mass Pop III stars is in dark green triangles, model PopIII(CCSNe+FSNe) including yields from only low-mass Pop III stars is in light green triangles, and model PISNe including yields from PISNe is in red diamonds. Error bars are computed following the method discussed in Sect. 3.6.

Table 3. Total number of supernova events, number of PISNe, and their corresponding iron ejection mass for the different IMFs used in this work.

Model ID	#SNe [1]/ #PISNe [2]	Fe Total [3]/ Fe PISNe [4]	Fe Boost [5]
PopII	7.4/0	0.29/0	1
PopIII(CCSNe)	33.2/0	2.6/0	9.0
PopIII(CCSNe+FSNe)	26.7/0	1.9/0	6.6
PopIII(CCSNe+FSNe)+PISNe	23.2/0.83	7.7/6.2	26.6

Notes. All quantities were computed for $1000\,M_\odot$ of stars formed. [1]Total number of supernovae (CCSNe and PISNe) for $1000\,M_\odot$ [2]Number of PISNe for $1000\,M_\odot$ of stars [3]Total iron mass ejected from CCSNe and PISNe per $1000\,M_\odot$ of stars [4]Iron mass ejected from PISNe per $1000\,M_\odot$ of stars [5]Ratio of the total mass of ejected iron in each model compared to the Pop II, which contains no Pop III stars.

galaxy models (see Table 2) considering only Pop II stars and checked the impact of our updates on the luminosity-metallicity relation. Our new galaxy models (PopII) are displayed in blue in Fig. 6. Compared to the old ones (shown also by blue circles in Fig. 1), the differences are small. With the new simulation setting, our Sextans model h177 ($L_V = 3\times10^5\,L_\odot$, [Fe/H] = -1.97), is perfectly superimposed with its counterpart simulated in Revaz & Jablonka (2018). At low luminosities, in the UFD regime, the strong deviation with respect to the observations remains. Our faintest dwarfs with luminosities of a few $10^3\,L_\odot$ have a [Fe/H] value as low as -3.5, about one dex below the observed relation. The new simulation setting slightly reshuffles the stellar luminosity and metallicity of our faintest dwarfs when studied individually, on a one to one basis. This underlines the great sensitivity of those objects to any kind of perturbations. Similar changes would have been expected by, for example, a modification to the random seed used in the stochastic star formation recipe (see e.g., Revaz et al. 2016; Keller et al. 2019).

From this first experiment, it stands out that using a higher resolution, but also an updated injection scheme and new yields for Pop II stars, do not solve the under-prediction of stellar metallicity at low luminosity. This first set of simulations with only Pop II stars is considered as our reference model PopII (see Table 1).

4.3. Pop III stars – without PISNe

Looking for the impact of Pop III stars in the evolution of the UFDs, we first considered them up to $140\,M_\odot$ (i.e., excluding PISNe). Pop III stars with initial masses between 13 and $30\,M_\odot$ explode as CCSNe. For stars with masses between 30 to $140\,M_\odot$, we considered two possibilities: (i) these stars are FSNe (i.e. the do not release any metal nor energy). This is

the PopIII(CCSNe+FSNe) model in Table 1 (Kobayashi et al. 2020). (ii) they explode as normal CCSNe with the nucleosynthesis yields of N13. Even if unlikely, this hypothesis was considered to check how much the released mass of metals can be increased in this stellar mass range. This is the PopIII(CCSNe) model in Table 1.

The impact of Pop III on the luminosity-metallicity relation is presented in Fig. 6. Light- and dark-green triangles stand for the PopIII(CCSNe) and PopIII(CCSNe+FSNe) models, respectively. Two effects explains the shift of the relation: on the one hand, Pop III stars explode in a few Myrs, and therefore do not contribute to the final stellar mass budget of their host galaxy. Compared to the PopII only model, the total luminosity is reduced by ∼30%. On the other hand, the IMF of Pop III stars starts at $13\,M_\odot$, strongly increasing the number of supernovae formed per unit of stellar mass. They can all potentially contribute to the metal enrichment. In contrast, in the case of the PopII model, a large fraction is locked in low-mass stars ($[0.05-8]\,M_\odot$), which do not contribute to the chemical enrichment of their host galaxy.

To illustrate this point, Table 3 provides the number supernovae that explode for $1000\,M_\odot$ of stars formed, as well as the amount of iron produced by each type of supernovae. While on average only seven CCSNe are present in the PopII model, this number rises to 26 for model PopIII(CCSNe+FSNe) and to 33 for PopIII(CCSNe). This increases in the number of supernovae boosts the released mass of iron. It is multiplied by 6.6 in PopIII(CCSNe+FSNe) and by nine in PopIII(CCSNe), as compared to PopII.

While the tension with observations is mitigated, thanks to the decrease in galaxy luminosity and the increase in metallicity, [Fe/H] still deviates by one dex, for luminosities below $5\times10^4\,L_\odot$.

4.4. Pop III stars – including PISNe

We next introduced very massive Pop III stars in our simulations, which explode as PISNe. To do so, we extended the maximal mass of the Pop III IMF from 140 to $300\,M_\odot$. The impact of PISNe on the luminosity-metallicity relation is shown by the red diamonds in Fig. 6. While there are still few faint galaxies that lie

slightly below the observed relation, with some having [Fe/H] ≤ −3, there is a general improvement, arising from: (i) an increase in metallicity, and (ii) a decrease in luminosity.

First, as seen in the top panel of Fig. 2, stars with masses above 140 $M_{\odot}$ provide the largest masses of iron (up to 100 times higher than the normal CCSNe). Table 3 shows that an IMF including PISNe generates 26 times more Fe than in `PopII`, and three times more than in `PopIII(CCSNe)`, with PISNe contributing to about 80% of this amount. Second, because Pop III stars dominate the UFD population but do not contribute to the total final mass budget, the luminosity of the `PISNe` model is again 0.2 dex lower than that of `PopII`, similar to the `Pop III` models without PISNe. The higher the fraction of Pop III relative to Pop II stars, the more these two effects impact the final luminosity and metallicity of the galaxy. They are particularly significant for the faintest dwarf galaxies, which have shorter star formation histories, as seen in Fig. 6. However, as PISNe may be extremely rare, even nonexistent in systems forming few stars, the metallicity is not increased at the faintest end (i.e., at a luminosity of about $10^3 L_{\odot}$). Figure 7 quantifies the origin of the iron content of each stellar particle in the `PISNe` model. In order to do so, the gas particles from which each of the stellar particle form were identified and their pollution history was analyzed. This was done thanks to a logger module that records on-the-fly the time at which a gas particle received metals from a supernova, allowing us to distinguish between CCSNe and PISNe sources.

Figure 7 shows the iron mass fraction originating from PISNe as a function of galaxy luminosity. The color code shows the relative number of PISNe. The trend is very clear. In the faintest systems, up to 4% of supernovae are PISNe, while the same value drops below 1% for the brightest systems. This means that the iron production by PISNe is dominant (up to 90%) in the total iron budget of the faint galaxies, while this value drops down to 30% for a Sextans-like galaxy. Indeed, Sextans and more massive dwarf galaxies are dominated by in situ star formation. Thus, after a short period of Pop III star formation, the galaxy is dominated by Pop II stars (see Fig. 11). On the contrary, in UFDs, as mentioned in Sect. 4.1, the first stars form in isolated small mini-halos. Due to their shallow gas reservoir, only a few stars can form in the same mini-halo. The gas metallicity stays below the metallicity threshold for the onset of the formation of Pop II stars up to 1 Gyr. As a consequence, Pop III, and in particular PISNe, have a strong impact on the metallicity and luminosity of the faintest UFDs.

There are two notable exceptions to this general rule: the halo `h291` with $L_V \cong 7 \times 10^2 L_{\odot}$ does not experience any PISNe explosion, and `h273` with $L_V \cong 2 \times 10^3 L_{\odot}$ has only one PISNe exploding in an isolated sub-halo. Indeed, for UFDs with luminosities below $10^4 L_{\odot}$, the expected number of PISNe is very low (see Table 3). This is confirmed by our models that, on average, form fewer than five PISNe per UFD. Thus, due to this stochasticity, PISNe induce a large scatter in the luminosity-metallicity relation rather than a shift toward a higher [Fe/H]. Consequently, PISNe cannot be counted as a reliable source of metals in these systems. This also underlines the importance of examining a large sample of UFDs, as done in this study.

4.5. Abundance ratios

Our simulations allowed us to assess the impact of the different types of first stars on the galaxy mean properties (e.g., [Fe/H], L_V), as well as their influence on the chemical abundance ratios of long-lived low-mass stars. This was done in comparison with the observations, compiling all existing chemical
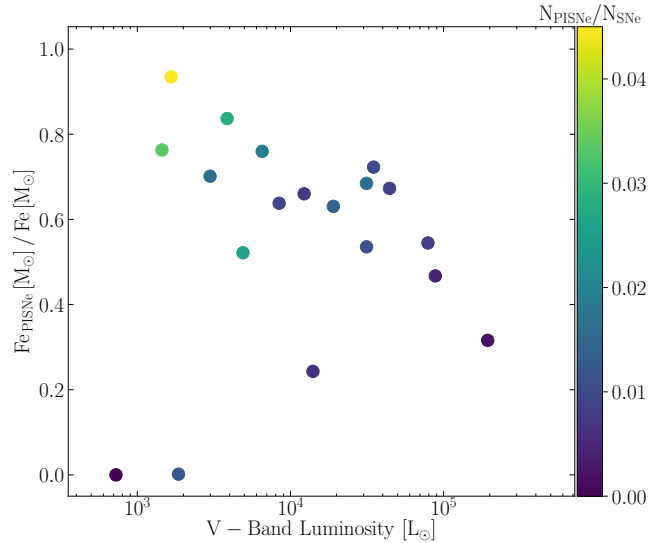

Fig. 7. Fe released by PISNe compared to the total Fe produced in all CCSNe and PISNe events. Each point represents one model galaxy, color-coded by the ratio between the number of stars forming PISNe and the total number of stars ending up as supernovae.

abundance studies with available abundance ratios from high or medium spectroscopic resolution.

Figure 8 presents the [Mg/Fe] and [Ca/Fe] stellar abundance ratios as a function of [Fe/H] for our Sextans model (`h177`) and different prescriptions listed in Table 1. These two α-elements are particularly well documented for the Sextans dSph, handling both nucleosynthesis and star formation timescales. In Figs. 9 and 10, for our UFD models, we restricted the comparison with observations to Mg. This is indeed the element that has been most often determined in these faint systems. However, only a handful of stars in each individual UFD benefit from detailed chemical analysis (the maximum being Bootes with 13 stars). Therefore, we drew the comparison of our 18 UFD models with the combination of the existing observations of 68 stars in 14 UFDs. Both simulations and observations were, however, split in two luminosity intervals, with $L_V > 10^4 L_{\odot}$ and $L_V < 10^4 L_{\odot}$. In Figs. 8–10, the model probability distributions are shown in blue, while the observations are displayed with gray points. The black and red contours delineate the regions encompassing 30, 60, and 90% of the stars, in the models and observations, respectively. In Fig. 8 it is indeed puzzling that a clump of stars exists at [Fe/H] $\cong$ −3. However, with the sparsity of data in this low metallicity range, it is not clear if this clump has a physical origin. On the contrary, our models containing a higher number of stars lead to a more continuous distribution.

All models of the `h177` halo feature the same general trend with increasing metallicity, starting with a super solar [α/Fe] plateau below [Fe/H] $\sim$ −2, followed by a knee and a descending branch down to subsolar [α/Fe] values, under the influence of the iron-rich Type Ia supernova ejecta. This is the consequence of a relatively extended star formation history, over 3 Gyr, imprinted in the Pop II stars, which constitute up to 98% of the total stellar populations in these massive systems (as shown in Fig. 11). However, some differences between the models are noticeable, which stem from the specific nucleosynthesis of each type of Pop III stars, but also the total number of supernova explosions and feedback energy. In particular, keeping as reference the `PopII` model, the highest concentration of stellar particles is located at lower [Fe/H] values, by about 0.2 dex for

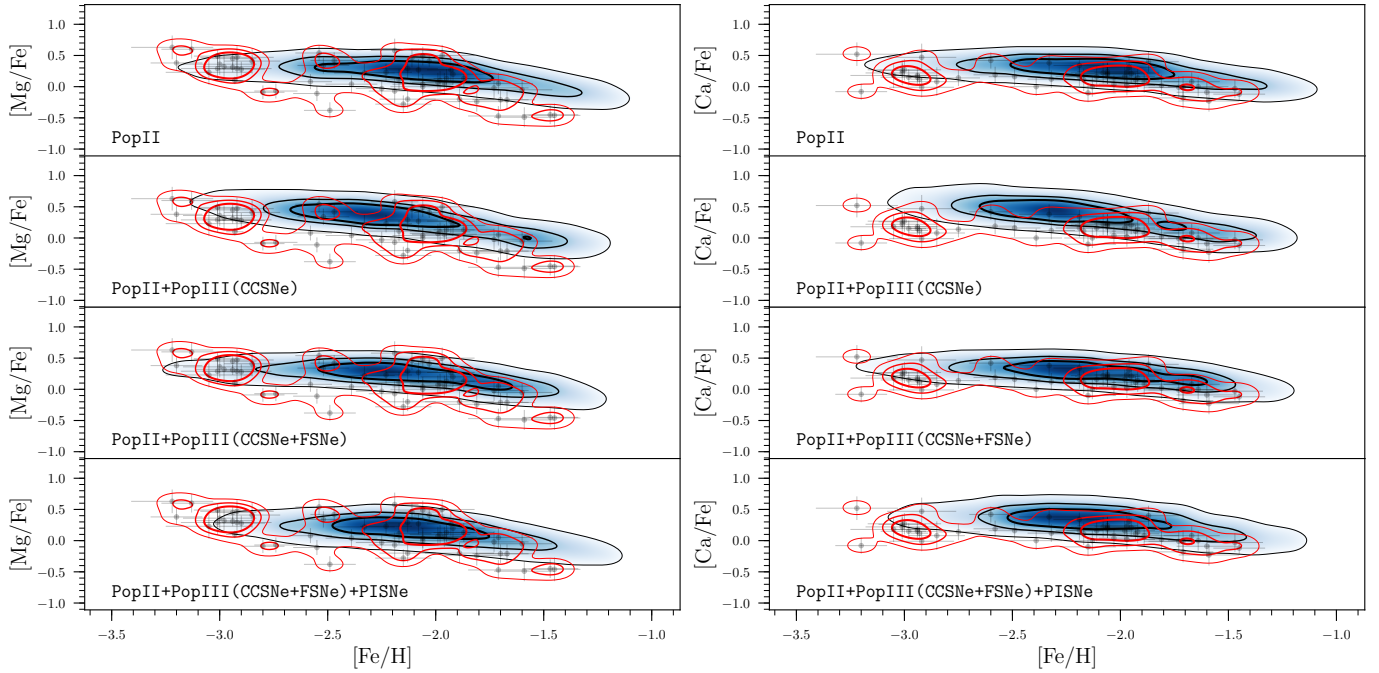**Fig. 8.** [Mg/Fe] and [Ca/Fe] abundance ratios as a function of [Fe/H] for the dwarf spheroidal model (h177). The stellar distribution in blue is compared to the Sextans observations, shown as gray points with error bars. For both simulations and observations, we added contours, shown in black and red, respectively, which represent the regions encompassing 30, 60, and 90% of the stars. The observational data are the same as in Fig. 5.

the PopIII(CCSNe) model. Moreover, a small but still significant fraction of the population has higher [Ca/Fe] than in the observations at [Fe/H] < −2, suggesting that it is indeed unlikely that stars in the mass range 30–140 $M_\odot$ contributed to the galaxy chemical enrichment, in agreement with the FSNe mass range of Kobayashi et al. (2020).

Figures 9 and 10 enable a deeper assessment of the impact of the Pop III stars in the evolution of UFDs above and below $10^4 L_\odot$, respectively. In both luminosity ranges, the main change in the chemical evolution of UFDs is the shift of the metallicity peak toward higher values, as seen in Fig. 6. For the brightest UFDs, the tail of the distribution of stars at [Fe/H] ≤ −3.5 also disappears. Pop III stars remain influential in these bright UFDs, still accounting for 38% of all stars formed (see Fig. 11), but the choice of the type of Pop III has little influence on the position of the [Fe/H] peak and the distribution in both [Mg/Fe] and [Fe/H] around it. Indeed, in these systems, the global galaxy evolution is dominated by the Pop II stars. This explains the limited sensitivity to the specific nature (CCSNe, FSNe or PISNe) and mass range of the first stars. However, PISNe, which on average produce less magnesium than iron, induce a decrease in [Mg/Fe] and the formation of metal-rich α-poor stars, accentuating the subsolar [Mg/Fe] branch to a level that does not appear to be supported by the observations.

The faintest UFDs, for which Pop III stars contribute to 86% of the full population (see Fig. 11), are more sensitive to the nature of the first stars, with the [Fe/H] distribution becoming slightly more peaked. The tail of the [Fe/H] distribution below −3.5 remains, irrespective of the Pop III nucleosynthesis prescription. More importantly, the intrinsic chemical evolution of these systems hardly exceeds [Fe/H] ∼ −2.5, contrary to the observations, which seem to extend much beyond. Given the extremely short star formation history of these systems, it is very difficult to see how their intrinsic evolution alone could ever reach metallicities as high as −1. Further comparison with the
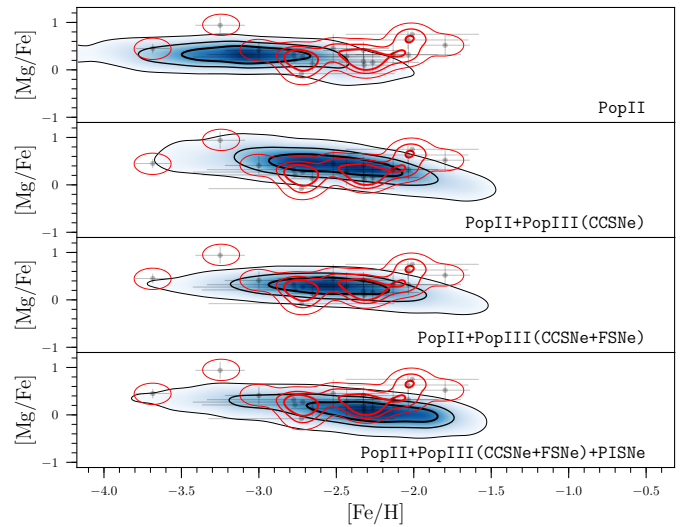

Fig. 9. [Mg/Fe] vs. [Fe/H] for the UFDs with luminosities larger than $10^4 L_\odot$, for the models (in blue) and the observations (in gray). Contours, in black for the models and in red for the observations, represent the regions encompassing 30, 60, and 90% of the stars. The data contain a total of 19 stars from Bootes I (Gilmore et al. 2013; Feltzing et al. 2009; Norris et al. 2010b; Ishigaki et al. 2014) and Hercules (Koch et al. 2008; Vargas et al. 2013; François et al. 2016).

observations, based on the stellar metallicity distribution, is conducted in Sect. 4.6.3.

4.6. Robustness of the results

We checked the robustness of our results with respect to different model parameters that could potentially affect our conclusions, specifically the gas metallicity threshold [Fe/H]$_c$ and the
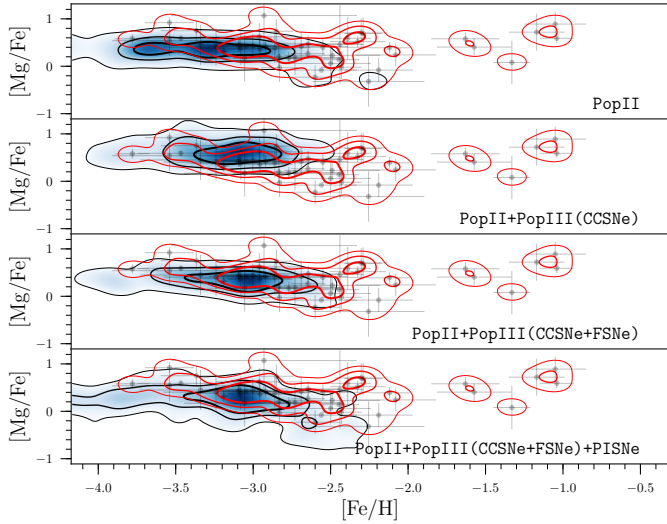
Fig. 10. [Mg/Fe] vs. [Fe/H] for the UFDs with luminosities lower than $10^4 L_\odot$, for the models (in blue) and the observations (in gray). Contours, in black for the models and in red for the observations, represent the regions encompassing 30, 60, and 90% of the stars. The data contain a total of 49 stars from Bootes II (Ji et al. 2016d; Koch et al. 2009; François et al. 2016), Canes Venatici II (François et al. 2016; Vargas et al. 2013), Grus I (Ji et al. 2019), Horologium (Nagasawa et al. 2018), Leo IV (Simon et al. 2010; François et al. 2016), Reticulum II (Ji et al. 2016c,a; Roederer et al. 2016; Ji & Frebel 2018), Segue I (Norris et al. 2010a; Vargas et al. 2013; Frebel et al. 2014), Segue II (Kirby et al. 2013), Triangulum II (Venn et al. 2017; Kirby et al. 2017; Ji et al. 2019), Tucana II (Ji et al. 2016b; Chiti et al. 2018), Tucana III (Hansen et al. 2017), and Ursa Major II (Frebel et al. 2010; Vargas et al. 2013).
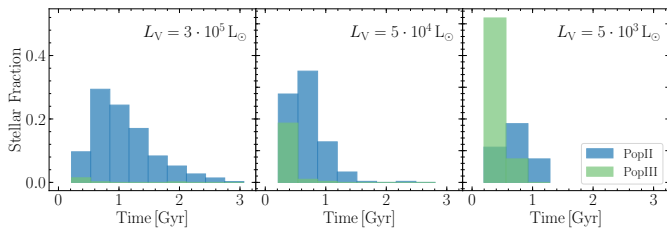
Fig. 11. Fraction of formed Pop II (in blue) and Pop III (in green) stars. Each panel corresponds to one halo, with its total luminosity indicated on the upper right side of the panel. The left panel shows our Sextans-like model, h177, with $L_V \sim 3\times10^5 L_\odot$. The middle panel shows h152 ($L_V \sim 0.5\times10^5 L_\odot$), representative of medium-mass UFDs. The right panel displays h323, representative of low-mass UFDs ($L_V \sim 0.05\times10^5 L_\odot$), with quenched star formation after 1 Gyr.

feedback energy of PISNe, which are uncertain. Then we looked into the comparison of the observed and model stellar metallicity distributions, in order to help provide deeper insight into the resisting shift between their corresponding [Fe/H].

4.6.1. Metallicity threshold [Fe/H]$_c$

We re-simulated the PISNe model, increasing [Fe/H]$_c$ from the fiducial value of -5 to -4. Increasing [Fe/H]$_c$ leads to an increased number of Pop III stars and consequently a potential increase in the final metallicity of UFDs, given that Pop III stars can provide up to three times more iron than Pop II stars (Sect. 4.3). Figure 12 compares the final metallicity of the [Fe/H]$_c = -5$ (red diamonds) and [Fe/H]$_c = -4$ (blue trian-

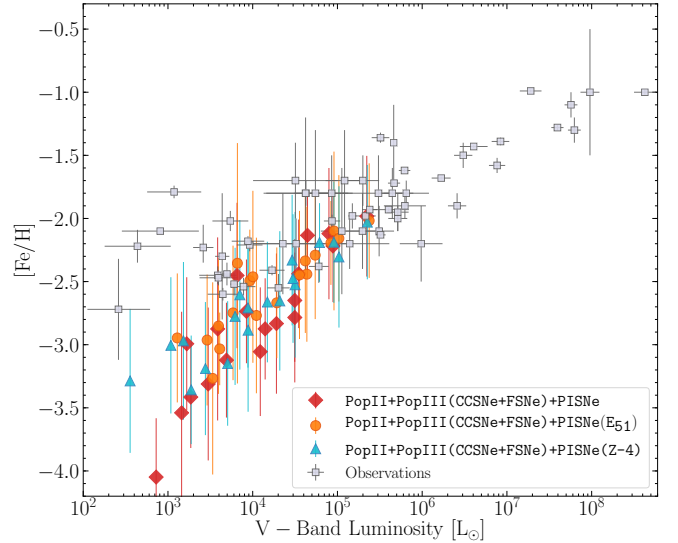**Fig. 12.** Tests of different prescriptions. The PISNe model is shown by the same red diamonds as in Fig. 6. Models with a fixed value of 1×10^{51} erg are shown as orange circles. Models with a critital metallicity of -4 are displayed with blue triangles.

gles) models as a function of metallicity. For the vast majority of the models, the impact was very limited. It was the strongest for the faintest systems such as h291 and h273, because the probability of PISNe explosions is very low in these low-mass systems (Sect. 4.4), and the addition of even one single event is very significant. Nevertheless, increasing [Fe/H]$_c$ does not lead to the monotonic upward shift of the global L_V-[Fe/H]$_c$ relation, which would make the models and observations overlap.

4.6.2. PISNe energy

As introduced in Sect. 4.4, the energy of PISNe strongly increases with the progenitor mass, from 9 to 5×10^{51} erg (see Fig. 2). Between zero and five PISNe can explode in faint UFDs ($L_V < 10^4 L_\odot$), and up to 15 in brighter ones. Thus, in extreme cases, more than 10^{53} erg can be released in the ISM, limiting the star formation activity in the early stages of the galaxy evolution.

We re-simulated the PISNe model, decreasing the energy of PISNe explosions to a constant value of 10^{51} erg, corresponding to the standard CCSNe feedback energy. In Fig. 12 the orange circles represent the PISNe(E$_{51}$) model with 10^{51} erg feedback, while the red diamonds show our fiducial PISNe model. The net effect of the reduction of the feedback energy is that there is no system below $10^3 L_\odot$ formed anymore, most likely because star formation was less hampered by the PISNe explosions. Similar to increasing [Fe/H]$_c$, while the metallicities slightly increased at a fixed luminosity, this was not in a sufficiently large amount to make the observed and modeled relations match.

4.6.3. UFD metallicity distribution functions

The challenge of reproducing the UFDs called for an even closer look at the structure of the models, but also at the status of the observations and in particular at how well they sample the properties of these galaxies. To this end, Fig. 13 compares the metallicity distribution function (MDF) of the PISNe(Z-4) model, chosen because it is the closest to the observed metallicity-luminosity relation in Fig.12, with the observed distributions derived from the calcium triplet, in three luminosity intervals.
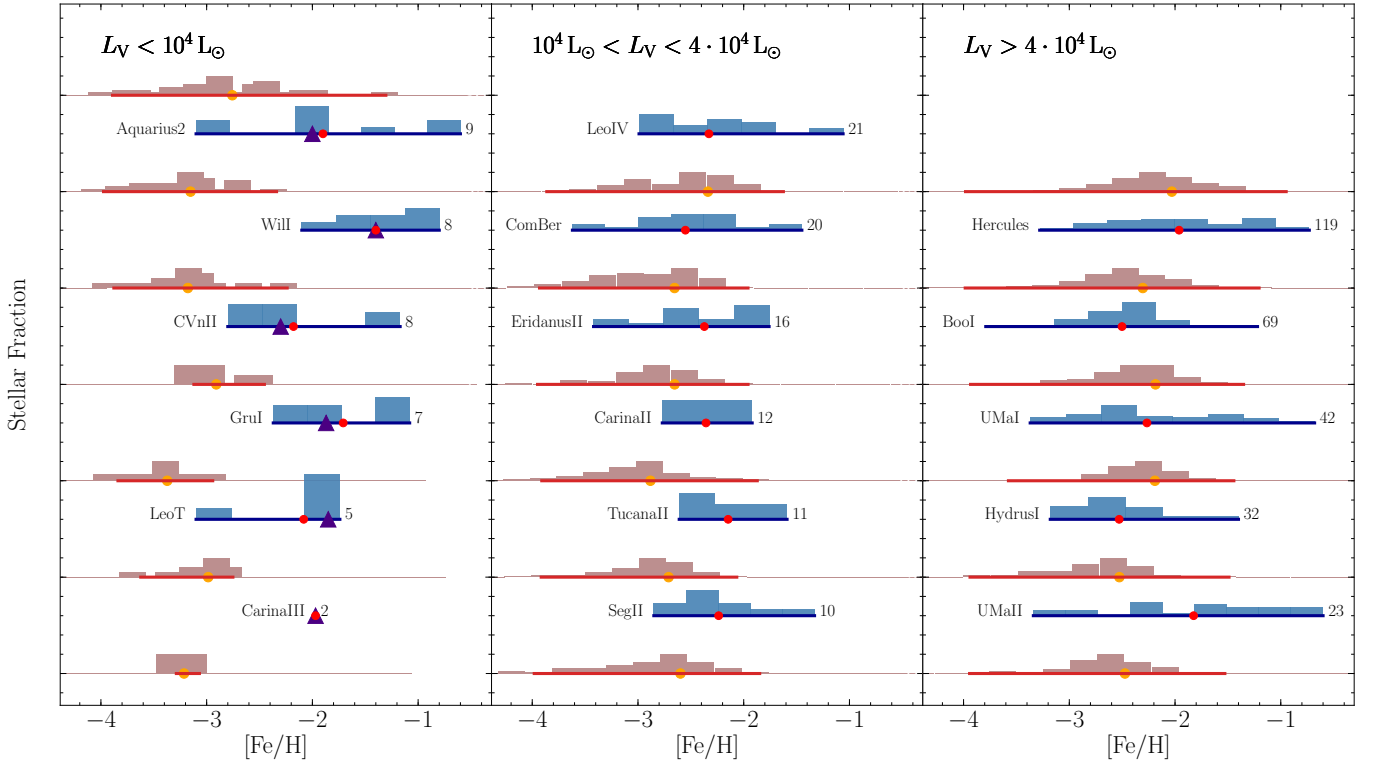
Fig. 13. Comparison of the metallicity distribution functions of the `PISNe(Z-4)` model (in red) and a set of observed galaxies (in blue) in three luminosity ranges. The solid red and blue horizontal lines indicate the full extent of the stellar metallicity distributions in the simulations and observations, respectively. The orange points show the mode of the [Fe/H] model distribution (see Appendix A). The purple triangles and red points represent, respectively, the median and mean of metallicities for each observed galaxy. The number of the spectroscopically confirmed member stars with metallicity estimates is indicated.

The stellar samples were taken from the references in Battaglia et al. (2022) and McConnachie (2012).

The observed MDF of each galaxy is displayed in blue. The number of spectroscopically confirmed galaxy members used in these MDFs is provided. The blue horizontal lines indicate the full metallicity ranges. The red points locate the mean [Fe/H] values. For the faintest systems, with fewer than ten member stars, we indicate the median [Fe/H] value with purple triangles. Although the number of member stars with available spectroscopic metallicities is fewer than 20 in most of the UFDs, it is already clear that they show a large variety of properties and a wide spread in metallicities up to more than 2 dex. The MDFs of the simulated galaxies are shown in red and their full metallicity ranges are indicated with solid red lines. The MDF peaks are shown with orange points. They are calculated in Appendix A.

As expected, the number of spectroscopically confirmed members decreases with the luminosity of the galaxies, with the consequence that the metallicity distributions may not be fully sampled. They are, in any case, very fragmented, making the position of their mean values rather uncertain. Less expected is that the low-metallicity tail, at [Fe/H] ≤ −3, seems to gradually disappear toward less massive systems. Conversely, the high-metallicity tail of the distributions remains. The question arises as to whether this apparent bias is simply a reflection of incomplete sampling or of different intrinsic evolution. In contrast, by their very nature, the models include all stellar populations. They keep the majority of the metal-poor population regardless of the mass and/or luminosity of the galaxy, whereas the full range of metallicity covered depends on the extent of the star formation history. As commented in the previous section, there

are few stellar particles above [Fe/H] = −2, and only in one case do we reach [Fe/H] = −1. In summary, it appears from this analysis that we are facing two limits, one observational, with the sampling of UFD properties still partial, but also a numerical limit with the challenge of reproducing systems that, despite very short star formation histories, contain stars as chemically enriched as in more massive systems.

5. Conclusions and discussion

Despite their many successes in recent years, the cosmological simulations of dwarf galaxies still fail at the smallest scales. Specifically, in the UFD regime, hydrodynamical models predict a steeper metallicity-luminosity relation than the one observed. Below $10^5\,L_\odot$, the galaxy mean metallicities can be as much as 2 dex too low at a given luminosity. In order to tackle this issue and understand how very faint galaxies acquire their metals, we studied the impact of Pop III stars on the metallicity evolution of UFDs. To this end, we performed cosmological zoom-in chemo-dynamical simulations from redshift 70 to zero of 19 halos hosting dwarfs with V-band luminosities between 2×10^3 to $3\times10^5\,L_\odot$. The most massive of these halos is a Sextans-like spheroidal dwarf galaxy, while the eighteen others are fully in the UFD regime, with luminosities equal or below $10^5\,L_\odot$. We validated our models by reproducing both the global properties and the α-elements trend of the Sextans dwarf galaxy. We checked the robustness of our results with respect to model-dependent parameters or poorly constrained physics. In particular, we looked into the effect of the critical metallicity, $[Fe/H]_c$, which defines the transition from Pop III to Pop II stars. We also

looked into the energy release from PISNe to see how it could affect the UFD build-up and physical properties.

Our main conclusions can be summarized as follows:

– In a ΛCDM cosmological framework, dwarf galaxies, and particularly UFDs, are dominated by metal-poor stars, which form in clumps hosted by the initial mini-halos. After a few hundred Myrs, these clumps merge and form the final galaxy.

– The evolution of UFDs is strongly influenced by Pop III stars. On average, 38% of all stars formed from gas at [Fe/H] < -5. In small mass UFDs, for which star formation is completely quenched after 1 Gyr, this fraction rises to 86%.

– Assuming a Kroupa-like IMF for Pop III stars, one with a slope of -1.3, as is usually considered for the massive stars, the main change in the chemical evolution of the dwarf galaxies due to the contribution of the first stars' nucleosynthesis is the upward shift of the peak of their final metallicity distributions. The nature of these first stars, which end up as CCSNe, FSNe or PSINe, is of secondary importance.

– Pop III stars with masses lower than $140\,M_\odot$ can increase the production of iron by up to a factor of six per $1000\,M_\odot$ of stars formed. This increase is, however, still insufficient to match the observed metallicity-luminosity relation.

– PISNe, with progenitor masses between $140\,M_\odot$ and $300\,M_\odot$, increase the production of iron by a factor of 26 per $1000\,M_\odot$ of stars formed. However, PISNe are rare events and occasionally absent in the faintest UFDs. Therefore, they have a limited impact on the global L_V – [Fe/H] relation.

The analysis of the galaxy metallicity distributions and trends of the stellar particle abundance ratios as a function of [Fe/H] provides further insight into the impact of Pop III stars:

– For the brightest UFDs, above $L_V = 10^4\,L_\odot$, the tail of the distribution of stars at [Fe/H] ≤ -3.5 disappears, while it remains for the fainter ones. None of the $L_V \geq 10^4\,L_\odot$ model UFDs contain stellar particles above [Fe/H] $= -2.5$, contrary to the observations. It remains a clear challenge to produce very small mass systems whose intrinsic evolution can create stars up to [Fe/H] $= -1$ on timescales smaller than $\sim$1 Gyr.

– Because PISNe release less magnesium than iron, low-mass long-lived stars formed from gas enriched by their ejecta can have subsolar [Mg/Fe]. These seems to be outnumbered at [Fe/H]≥ -2.5 compared to the observations of UFDs.

– The analysis of the model abundance ratios rules out any significant contribution from Pop III stars in the $30\,M_\odot$ to $140\,M_\odot$ mass range. Indeed, they predict, for all dwarf galaxies, supersolar [Ca/Fe] and [Mg/Fe] ratios, in contradiction with the observations.

– From an observational point of view, the number of known, spectroscopically confirmed members, with metallicity estimates and/or detailed chemical abundances, is still limited for all UFDs. It decreases even more with galaxy luminosity. As a consequence, the observed stellar metallicity distributions can be still very fragmented, making the position of their mean values uncertain. Moreover the low-metallicity tail, at [Fe/H] ≤ -3, seems to gradually disappear toward less massive systems.

Challenges to be taken up are thus both observational and numerical, with the need to reproduce systems that, despite very short star formation histories, contain stars as chemically enriched as in more massive ones.

References

Agertz, O., Pontzen, A., Read, J. I., et al. 2020, MNRAS, 491, 1656
Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
Aoyama, S., Hirashita, H., Lim, C.-F., et al. 2019, MNRAS, 484, 1852
Applebaum, E., Brooks, A. M., Christensen, C. R., et al. 2021, ApJ, 906, 96
Aubert, D., & Teyssier, R. 2010, ApJ, 724, 244
Barkat, Z., Rakavy, G., & Sack, N. 1967, Phys. Rev. Lett., 18, 379
Bate, M. R., & Burkert, A. 1997, MNRAS, 288, 1060
Battaglia, G., Tolstoy, E., Helmi, A., et al. 2011, MNRAS, 411, 1013
Battaglia, G., Taibi, S., Thomas, G. F., & Fritz, T. K. 2022, A&A, 657, A54
Bekki, K. 2015, MNRAS, 449, 1625
Bettinelli, M., Hidalgo, S. L., Cassisi, S., Aparicio, A., & Piotto, G. 2018, MNRAS, 476, 71
Bovill, M. S., & Ricotti, M. 2009, ApJ, 693, 1859
Bovill, M. S., & Ricotti, M. 2011, ApJ, 741, 17
Bromm, V. 2013, Rep. Prog. Phys., 76, 112901
Bromm, V., & Larson, R. B. 2004, ARA&A, 42, 79
Bromm, V., & Loeb, A. 2003, Nature, 425, 812
Bromm, V., Coppi, P. S., & Larson, R. B. 1999, ApJ, 527, L5
Brown, T. M., Tumlinson, J., Geha, M., et al. 2014, ApJ, 796, 91
Bullock, J. S., & Boylan-Kolchin, M. 2017, ARA&A, 55, 343
Chiaki, G., Wise, J. H., Marassi, S., et al. 2020, MNRAS, 497, 3149
Chiti, A., Frebel, A., Ji, A. P., et al. 2018, ApJ, 857, 74
Chon, S., Omukai, K., & Schneider, R. 2021, MNRAS, 508, 4175
Chon, S., Ono, H., Omukai, K., & Schneider, R. 2022, MNRAS, 514, 4639
Clark, P. C., Glover, S. C. O., Smith, R. J., et al. 2011, Science, 331, 1040
de Bennassuti, M., Salvadori, S., Schneider, R., Valiante, R., & Omukai, K. 2017, MNRAS, 465, 926
Dixon, K. L., Iliev, I. T., Gottlöber, S., et al. 2018, MNRAS, 477, 867
Escala, I., Wetzel, A., Kirby, E. N., et al. 2018, MNRAS, 474, 2194
Feltzing, S., Eriksson, K., Kleyna, J., & Wilkinson, M. I. 2009, A&A, 508, L1
François, P., Monaco, L., Bonifacio, P., et al. 2016, A&A, 588, A7
Frebel, A., Simon, J. D., Geha, M., & Willman, B. 2010, ApJ, 708, 560
Frebel, A., Simon, J. D., & Kirby, E. N. 2014, ApJ, 786, 74
Gallart, C., Monelli, M., Ruiz-Lara, T., et al. 2021, ApJ, 909, 192
Gandhi, P. J., Wetzel, A., Hopkins, P. F., et al. 2022, MNRAS, 516, 1941
Garnett, D. R. 2002, ApJ, 581, 1019
Gilmore, G., Norris, J. E., Monaco, L., et al. 2013, ApJ, 763, 61
Graziani, L., de Bennassuti, M., Schneider, R., Kawata, D., & Salvadori, S. 2017, MNRAS, 469, 1101
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Hahn, O., & Abel, T. 2011, MNRAS, 415, 2101
Hansen, T. T., Simon, J. D., Marshall, J. L., et al. 2017, ApJ, 838, 44
Harvey, D., Revaz, Y., Robertson, A., & Hausammann, L. 2018, MNRAS, 481, L89
Hausammann, L., Revaz, Y., & Jablonka, P. 2019, A&A, 624, A11
Heger, A., & Woosley, S. E. 2010, ApJ, 724, 341
Hicks, W. M., Wells, A., Norman, M. L., et al. 2021, ApJ, 909, 70
Hirashita, H. 2015, MNRAS, 447, 2937
Hopkins, P. F. 2013, MNRAS, 428, 2840
Hummel, J. A., Pawlik, A. H., Milosavljević, M., & Bromm, V. 2012, ApJ, 755, 72
Ishigaki, M. N., Aoki, W., Arimoto, N., & Okamoto, S. 2014, A&A, 562, A146
Iwamoto, N. 2005, Science, 309, 451
Jeon, M., Besla, G., & Bromm, V. 2017, ApJ, 848, 85
Jeon, M., Bromm, V., Besla, G., Yoon, J., & Choi, Y. 2021, MNRAS, 502, 1
Ji, A. P., & Frebel, A. 2018, ApJ, 856, 138
Ji, A. P., Frebel, A., & Bromm, V. 2015, MNRAS, 454, 659
Ji, A. P., Frebel, A., Simon, J. D., & Geha, M. 2016a, ApJ, 817, 41
Ji, A. P., Frebel, A., Simon, J. D., & Chiti, A. 2016b, ApJ, 830, 93
Ji, A. P., Frebel, A., Chiti, A., & Simon, J. D. 2016c, Nature, 531, 610
Ji, A. P., Frebel, A., Ezzeddine, R., & Casey, A. R. 2016d, ApJ, 832, L3
Ji, A. P., Simon, J. D., Frebel, A., Venn, K. A., & Hansen, T. T. 2019, ApJ, 870, 83
Karlsson, T., Johnson, J. L., & Bromm, V. 2008, ApJ, 679, 6
Katz, N. 1992, ApJ, 391, 502

Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19
Keller, B. W., Wadsley, J. W., Wang, L., & Kruijssen, J. M. D. 2019, MNRAS, 482, 2244
Kirby, E. N., Simon, J. D., Geha, M., Guhathakurta, P., & Frebel, A. 2008, ApJ, 685, L43
Kirby, E. N., Guhathakurta, P., Simon, J. D., et al. 2010, ApJS, 191, 352
Kirby, E. N., Boylan-Kolchin, M., Cohen, J. G., et al. 2013, ApJ, 770, 16
Kirby, E. N., Cohen, J. G., Simon, J. D., et al. 2017, ApJ, 838, 83
Knebe, A., Wagner, C., Knollmann, S., Diekershoff, T., & Krause, F. 2009, ApJ, 698, 266
Kobayashi, C., Tsujimoto, T., & Nomoto, K. 2000, ApJ, 539, 26
Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, ApJ, 900, 179
Koch, A., McWilliam, A., Grebel, E. K., Zucker, D. B., & Belokurov, V. 2008, ApJ, 688, L13
Koch, A., Wilkinson, M. I., Kleyna, J. T., et al. 2009, ApJ, 690, 453
Kodama, T., & Arimoto, N. 1997, A&A, 320, 41
Kroupa, P. 2001, MNRAS, 322, 231
Larson, R. B. 1998, MNRAS, 301, 569
Lee, M. G., Park, H. S., Park, J.-H., et al. 2003, AJ, 126, 2840
Lee, M. G., Yuk, I.-S., Park, H. S., Harris, J., & Zaritsky, D. 2009, ApJ, 703, 692
Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, A&A, 80, 155
Lucchesi, R., Lardo, C., Primas, F., et al. 2020, A&A, 644, A75
Macciò, A. V., Frings, J., Buck, T., et al. 2017, MNRAS, 472, 2356
Machida, M. N. 2008, ApJ, 682, L1
Maeder, A., & Meynet, G. 2012, Rev. Mod. Phys., 84, 25
Magg, M., Hartwig, T., Glover, S. C. O., Klessen, R. S., & Whalen, D. J. 2016, MNRAS, 462, 3591
Martin, N. F., Ibata, R. A., Chapman, S. C., Irwin, M., & Lewis, G. F. 2007, MNRAS, 380, 281
McConnachie, A. W. 2012, AJ, 144, 4
McKinnon, R., Torrey, P., Vogelsberger, M., Hayward, C. C., & Marinacci, F. 2017, MNRAS, 468, 1505
Muñoz, R. R., Carlin, J. L., Frinchaboy, P. M., et al. 2006, ApJ, 650, L51
Nagasawa, D. Q., Marshall, J. L., Li, T. S., et al. 2018, ApJ, 852, 99
Nichols, M., Revaz, Y., & Jablonka, P. 2014, A&A, 564, A112
Nichols, M., Revaz, Y., & Jablonka, P. 2015, A&A, 582, A23
Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, ARA&A, 51, 457
Norris, J. E., Yong, D., Gilmore, G., & Wyse, R. F. G. 2010a, ApJ, 711, 350
Norris, J. E., Wyse, R. F. G., Gilmore, G., et al. 2010b, ApJ, 723, 1632
Oñorbe, J., Garrison-Kimmel, S., Maller, A. H., et al. 2014, MNRAS, 437, 1894
Okamoto, T., Eke, V. R., Frenk, C. S., & Jenkins, A. 2005, MNRAS, 363, 1299
Okamoto, S., Arimoto, N., Yamada, Y., & Onodera, M. 2012, ApJ, 744, 96
Omukai, K., Tsuribe, T., Schneider, R., & Ferrara, A. 2005, ApJ, 626, 627
Ostriker, J. P., & McKee, C. F. 1988, Rev. Mod. Phys., 60, 1
Owen, J. M., & Villumsen, J. V. 1997, ApJ, 481, 1
Pagel, B. E. J. 1997, Nucleosynthesis and Chemical Evolution of Galaxies (UK: Cambridge University Press)
Peñarrubia, J., Navarro, J. F., & McConnachie, A. W. 2008, ApJ, 673, 226
Planck Collaboration XIII. 2016, A&A, 594, A13
Prgomet, M., Rey, M. P., Andersson, E. P., et al. 2022, MNRAS, 513, 2326
Price, D. J. 2012, J. Comput. Phys., 231, 759
Revaz, Y., & Jablonka, P. 2012, A&A, 538, A82
Revaz, Y., & Jablonka, P. 2018, A&A, 616, A96
Revaz, Y., Arnaudon, A., Nichols, M., & Jablonka, P. 2016, A&A, 588, A21
Ricotti, M., & Gnedin, N. Y. 2005, ApJ, 629, 259
Rodriguez Wimberly, M. K., Cooper, M. C., Fillingham, S. P., et al. 2019, MNRAS, 483, 4031
Roederer, I. U., Mateo, M., Bailey, John I., I., , et al. 2016, AJ, 151, 82
Sacchi, E., Richstein, H., Kallivayalil, N., et al. 2021, ApJ, 920, L19
Salvadori, S., & Ferrara, A. 2009, MNRAS, 395, L6
Sanati, M., Revaz, Y., Schober, J., Kunze, K. E., & Jablonka, P. 2020, A&A, 643, A54
Schauer, A. T. P., Glover, S. C. O., Klessen, R. S., & Clark, P. 2021, MNRAS, 507, 1775
Schmidt, M. 1959, ApJ, 129, 243
Sedov, L. I. 1959, Similarity and Dimensional Methods in Mechanics (New York: Academic Press)
Shetrone, M.-D., C-t, P., & Stetson, P.-B. 2001, PASP, 113, 1122
Simon, J. D. 2019, ARA&A, 57, 375
Simon, J. D., & Geha, M. 2007, ApJ, 670, 313
Simon, J. D., Frebel, A., McWilliam, A., Kirby, E. N., & Thompson, I. B. 2010, ApJ, 716, 446
Skillman, E. D., Kennicutt, R. C., & Hodge, P. W. 1989, ApJ, 347, 875
Skinner, D., & Wise, J. H. 2020, MNRAS, 492, 4386
Smartt, S. J. 2009, ARA&A, 47, 63
Smith, B. D., Bryan, G. L., Glover, S. C. O., et al. 2017, MNRAS, 466, 2217
Springel, V. 2005, MNRAS, 364, 1105
Springel, V. 2010, ARA&A, 48, 391
Springel, V., & Hernquist, L. 2002, MNRAS, 333, 649
Stacy, A., Greif, T. H., & Bromm, V. 2010, in First Stars and Galaxies: Challenges for the Next Decade, eds. D. J. Whalen, V. Bromm, & N. Yoshida, Am. Inst. Phys. Conf. Ser., 1294, 289
Stinson, G., Seth, A., Katz, N., et al. 2006, MNRAS, 373, 1074
Tafelmeyer, M., Jablonka, P., Hill, V., et al. 2010, A&A, 524, A58
Taylor, G. 1950, Proc. Roy. Soc. London Ser. A, 201, 159
Theler, R., Jablonka, P., Lucchesi, R., et al. 2020, A&A, 642, A176
Tornatore, L., Borgani, S., Dolag, K., & Matteucci, F. 2007, MNRAS, 382, 1050
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Truelove, J. K., Klein, R. I., McKee, C. F., et al. 1997, ApJ, 489, L179
Tsujimoto, T., Nomoto, K., Yoshii, Y., et al. 1995, MNRAS, 277, 945
Turk, M. J., Abel, T., & O'Shea, B. 2009, Science, 325, 601
Umeda, H., & Nomoto, K. 2002, ApJ, 565, 385
Vargas, L. C., Geha, M., Kirby, E. N., & Simon, J. D. 2013, ApJ, 767, 134
Vazdekis, A., Casuso, E., Peletier, R. F., & Beckman, J. E. 1996, ApJS, 106, 307
Venn, K. A., Starkenburg, E., Malo, L., Martin, N., & Laevens, B. P. M. 2017, MNRAS, 466, 3741
Wheeler, C., Hopkins, P. F., Pace, A. B., et al. 2019, MNRAS, 490, 4447
Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009, MNRAS, 399, 574
Wyithe, J. S. B., & Loeb, A. 2006, Nature, 441, 322
Yoshida, N., Hosokawa, T., & Omukai, K. 2012, Prog. Theor. Exp. Phys., 2012, 01A305

Appendix A: Determining [Fe/H] for simulated galaxies

Properly attributing a stellar metallicity to each simulated dwarf galaxy is essential in this work. Knowing the metallicity of each stellar particle, the dwarf metallicity is usually computed by taking the mean, the median, or the mode of the MDF. By taking the mean or the median, one would, however, strongly underestimate the system metallicity, owing to the presence of stars with very low metallicity in simulations that contain only Pop II stars. Taking the mode is more appropriate but this may be hardly determined in noisy MDFs. Increasing the number of metallicity bins will help, although at the expense of the precision.

We used a method where the galaxy stellar metallicity is derived from the mode of a fit to the MDF. The fitting function was chosen as the MDF predicted by a simple chemical evolution model that assumes an instantaneous recycling (see Pagel 1997) :

$$\frac{\mathrm{d}N}{\mathrm{d}[\mathrm{Fe/H}]} \propto 10^{[\mathrm{Fe/H}]} \exp\left(-\frac{10^{[\mathrm{Fe/H}]}}{p}\right). \tag{A.1}$$

Here, p is a free parameter that is related to the position of the MDF mode by the relation $[\mathrm{Fe/H}]_{\mathrm{mode}} = p \ln(10)$. Fig. A.1 illustrates the application of the method on three different dwarfs. It demonstrates its ability to derive a reliable mode, even if the histogram is noisy or if a large peak is found outside the main range of the MDF.
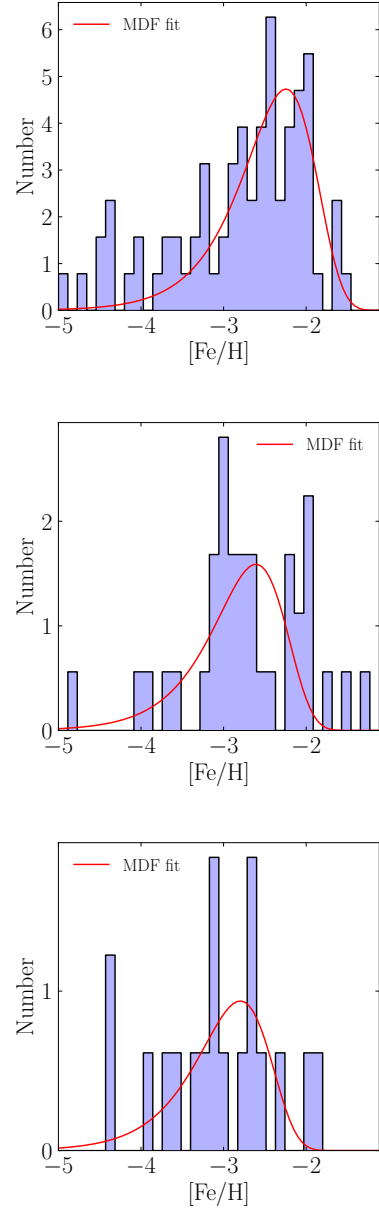

Fig. A.1. Application of the fitting method on three UFDs of decreasing luminosity (from left to right, $L_\mathrm{V} = 3.9 \cdot 10^4$ L$_\odot$, $L_\mathrm{V} = 1.9 \cdot 10^4$ L$_\odot$, and $L_\mathrm{V} = 1.0 \cdot 10^4$ L$_\odot$). The MDF fit given by Eq. A.1 is displayed in red.