Progenitor-mass-dependent yields amplify intrinsic scatter in dwarf-galaxy elemental abundance ratios

Dhruv A. Muley, Coral R. Wheeler, Philip F. Hopkins, Andrew Emerick, and Dušan Kereš

1 Department of Astronomy, University of California, Berkeley, Berkeley, CA 94720 USA
2 Department of Physics and Astronomy, University of Victoria, Victoria, BC V8W 2Y2, Canada
3 California Institute of Technology, Pasadena, CA 91125, USA
4 University of California, San Diego, La Jolla, CA 92093-0021, USA
5 Carnegie Observatories, Pasadena, CA 91101, USA
6 Department of Physics and Astronomy, University of California, Davis, Davis, CA 95616, USA

ABSTRACT
We explore the effect of including progenitor mass- and metallicity-dependent yields, supernova rates and energetics on variations in elemental abundance ratios (particularly [α/Fe]) in dwarf galaxies. To understand how the scatter and overall trends in [α/Fe] are affected by including variable metal yields from a discretely sampled initial mass function, we run FIRE simulations of a dwarf galaxy (M∗ = 0) using nucleosynthetic yields from the NuGrid data base that depend on the stellar progenitor mass and metallicity. While NuGrid exhibits lower aggregate α-element production than default FIRE yields, we find that its explicit mass dependence, even when including turbulent metal diffusion, substantially widens the intrinsic scatter in the simulated [Fe/H]-[α/Fe] – a phenomenon visible in some observations of dwarf galaxies.

Key words: galaxies: abundances – galaxies: dwarf – galaxies: evolution.

1 INTRODUCTION
Recent detailed observations of dwarf galaxies have helped constrain their elemental evolution and star formation histories (SFHs) with ever-increasing precision. Abundance measurements from medium-resolution spectroscopy can detect a variety of different species in galaxies, including Fe, Mg, Si, Ca, and Ti. These measurements have been used to estimate the overall stellar metallicity and metallicity distributions for Milky Way dwarf spheroidal (dSph) satellite galaxies (Shetrone, Côté & Sargent 2001; Cohen & Huang 2010; Tafelmeyer et al. 2010; Venn et al. 2012; Starkenburg et al. 2013), ultra-faint dSphs (Frebel et al. 2010; Vargas et al. 2013), and for satellites of M31 (Vargas, Geha & Tollerud 2014; Kirby et al. 2020; Wojno et al. 2020). These observations have revealed, among other features, a tight stellar mass–metallicity relation (MZR) spanning five orders of magnitude in stellar mass (Kirby et al. 2011a, 2013; Ho et al. 2014), narrow metallicity distributions (Kirby et al. 2010), and a generally decreasing trend in [α/Fe] versus [Fe/H] (Kirby et al. 2009, 2010; Tolstoy, Hill & Tosi 2009; Letarte et al. 2010; Hasselquist et al. 2017; Hill et al. 2019).

Comparing these high-quality observations to one-zone chemical evolution models (e.g. Lanfranchi & Matteucci 2003; Cescutti 2008; Kirby et al. 2011a,b) as well as hydrodynamical simulations, has improved our understanding of time-scales and yields for processes such as AGB winds, Type Ia supernovae, and Type II core-collapse supernovae, but some questions still remain. While many observational studies of dwarfs have found that they have a narrow scatter in α abundances (see e.g. Tolstoy et al. 2009; Revaz et al. 2016, for a compilation), observations of some dwarfs may indicate a much greater intrinsic scatter (Aoki et al. 2009; Lemasle et al. 2012), particularly at low metallicities/early times when Type II supernovae dominate enrichment (Frebel & Norris 2015; Vargas et al. 2013, 2014). The connection between progenitor mass, metallicity, and α-element yields from Type II supernovae has long been known (Tsujimoto et al. 1995; Woosley & Weaver 1995), and a number of simulated works have incorporated some form of progenitor mass- and metallicity-dependent yields into their chemical enrichment models (Mosconi et al. 2001; Kawata et al. 2006; Stinson et al. 2006; Wiersma et al. 2009; Schaye et al. 2015). However, few have done so with the resolution required to investigate dwarf galaxies – particularly with time resolution sufficient to allow for the mixing of the varying yields within a dynamical time (Few et al. 2012; Pillepich et al. 2018, but see Jeon, Besla & Bromm 2017; Revaz & Jablonka 2018 for high-resolution examples). Fewer still (see Emerick, Bryan & Mac Low 2019 for an exception) have specifically investigated how scatter in the [α/Fe] versus [Fe/H] relationship is affected by averaging yields from Type II supernovae and stellar winds (if included) over the initial mass function (IMF; Valcke, de Rijcke & Dejonghe 2008; Okamoto et al. 2010), as is still done in a number of simulations (Marcolini et al. 2008; Schroyen et al. 2013; Wang et al. 2015; Davé, Thompson & Hopkins 2016; Hirai & Saitoh 2017; Starkenburg et al. 2017; Davé et al. 2019). This includes the
standard FIRE-2 simulations from Hopkins et al. (2018) to which we directly compare in this work. One notable exception is Revaz et al. (2016), who use non-cosmological simulations of dwarf galaxies to investigate the effect that high resolution has on artificially increasing the scatter in $[\alpha/Fe]$ versus $[Fe/H]$ due to incomplete sampling of the IMF. They find that the scatter is reduced when a metal-mixing scheme is included, but do not explicitly compare to the case in which the yields are averaged over the IMF.

In this paper, we present a case-study of the effect of IMF averaging on $[\alpha/Fe]$ versus $[Fe/H]$ by re-simulating the $M_\star \sim 10^6 M_\odot$ m10q dwarf galaxy from the FIRE-2 project (Hopkins et al. 2018) with both IMF-averaged and properly progenitor mass- and metallicity-dependent yields. We modify the yield, SNe rate, and energy models in three different runs: the default, IMF-averaged, metallicity-independent FIRE-2 yields (m10q.FIRE-2, based on Leitherer et al. 1999; Nomoto et al. 2006; Hopkins et al. 2018), the progenitor-mass- (and metallicity-) dependent NuGrid data base of post-processed stellar evolution simulations (m10q.nugrid, based on Pignatari et al. 2016; Ritter et al. 2018a,b), and IMF-averaged NuGrid yields (m10q.nugrid_avg). We show that progenitor-mass dependence (but not metallicity-dependence alone) raises the scatter in $[\alpha/Fe]$-[Fe/H] to levels commensurate with some observations, but that neither changing the yields alone, nor adding in the mass and metallicity dependencies, have a dynamically significant effect on the simulated galaxy. However, we do find that, at low $[Fe/H]$, using NuGrid yields results in a systematically lower $[\alpha/Fe]$ abundance than is observed in many (but not all) dwarf galaxies, implying that its treatment of some yields (particularly Mg) may need to be revisited.

2 METHODS

In line with existing FIRE simulations, we use the meshless finite-mass (MFM) hydrodynamical method (Hopkins 2015) at a resolution of $m_{\text{bar}} = 250 M_\odot$ (e.g. Wheeler et al. 2019). We track 11 species: H, He, C, N, O, Ne, Mg, Si, S, Ca, and Fe; and three enrichment processes: Type II core-collapse supernovae, O/AGB winds, and Type Ia supernovae (Hopkins et al. 2018). Wind yields and energies are deposited continuously, while supernovae are drawn each timestep from a binomial distribution with probability $p$ based on the event rate, time-step duration, and star particle mass; when an event occurs, we assign its yields and energy as a function of star-particle age. For m10q.nugrid, we use the SYGMA (Ritter et al. 2018a) code to construct time- and metallicity-dependent supernova yields and event rates for a Kroupa (2002) stellar population, whereas for m10q.nugrid_avg, we average these yields over the IMF. In all cases, the IMF runs from 0.1–100 $M_\odot$, with a Type II supernova progenitor minimum mass of 8 $M_\odot$. As a check of the effect of this lower mass cut-off, we created a sample stellar population with an assumed 5–100 $M_\odot$ range of supernova progenitors, and found that the minimum and maximum $[\alpha/Fe]$ values in supernova ejecta remain the same, and so will not significantly affect the results presented here. On average, NuGrid core-collapse supernovae produce $\sim 13 M_\odot$ of ejecta per event, while the default FIRE-2 core-collapse supernovae produce $10.5 M_\odot$. Given the time-steps we use, $p$ never exceeds 1, so our supernovae are ‘time-resolved’ (i.e. we avoid over- or undercounting). This in turn ensures that on average, m10q.nugrid accurately reproduces the yield profile from the tables.

Because NuGrid’s AGB yields assume solar abundance ratios, we use them only for the total mass-loss rate, but then partition the ejecta between elements according to their surface abundances (as in Hopkins et al. 2018). We employ the same treatment in m10q.nugrid_avg after IMF-averaging to eliminate time dependence. For type Ia supernovae in both the m10q.nugrid and m10q.nugrid_avg simulations, we continue to use the pre-existing Hopkins et al. (2018) yields, but with a simple $R_{\text{type}} \propto t^{-1.1}$ delay-time distribution (from Maoz & Graur 2017) rather than the Gaussian prompt-constant delay time treatment (Mannucci, Della Valle & Panagia 2006) used in m10q.FIRE-2 and in previous FIRE-2 simulations. This increases the number of Type Ias by $\sim 2 \times$, to $\sim 15$ per cent of all supernovae, but does not otherwise affect the main results presented here.

In the top row of Fig. 1, we plot the rates, energies, and ejecta mass for core-collapse supernova events in NuGrid, alongside the STARBURST99-based (Leitherer et al. 1999) FIRE-2 default fits. As expected, the quantities obtained from NuGrid decline with time, due to the decreasing progenitor mass, with only a weak dependence on metallicity. In an integrated sense, however, the time-integrated difference in supernovae per solar mass formed between the two models is $\sim 20$ per cent, and differences in rate at any given time are no more than a factor of several. This implies (and our simulations bear out) that for galaxy-averaged properties such as SFRs and total metallicities, there should be no qualitatively meaningful difference between yield models.

In the second and third rows, we plot the ejecta mass fraction of various elements as a function of age (equivalently, progenitor mass) from the NuGrid tables, along with the IMF-averaged default FIRE-2 yields from Nomoto et al. (2006). For O, we find a high mass-fraction in ejecta from the most massive progenitors, followed by a metallicity-dependent drop-off; for N, production scales linearly in metallicity ($f_\text{N} \approx 4.05 \times 10^{-3}$ ($Z/Z_\odot$)), with the exception of massive progenitors at the lowest metallicity $Z = 10^{-4}$. Of the four $\alpha$-elements comprising the Kirby et al. (2011a) observational definition, we directly track Mg, Si, and Ca, plotted in the bottom row of Fig. 1. Mass fractions for these elements, along with Fe, exhibit sharp jumps at high progenitor mass, before stabilizing at later times. The times and magnitudes of these jumps differ for different metallicities (and the metallicity-dependence differs for each element), so the resulting $[\text{Mg}, \text{Si}, \text{Ca}/\text{Fe}]$ of a star particle’s core-collapse ejecta can vary by several dex between different core-collapse events from the same star particle. To facilitate comparison with default FIRE-2, we do not directly simulate Ti, but rather use Ca as a tracer in post-processing (estimating $[\alpha/Fe] = (1/4)[\text{Mg}/\text{Fe}] + [\text{Si}/\text{Fe}] + [\text{Ca}/\text{Fe}] + [\text{Ti}/\text{Fe}] \approx (1/4)[(\text{Mg}/\text{Fe}) + (\text{Si}/\text{Fe}) + 2(\text{Ca}/\text{Fe})]$). We motivate this choice by the fact that in the NuGrid tables, $[\text{Si}/\text{Fe}]$ for core-collapse ejecta almost always differs by $\lesssim 0.04$ dex between definition and estimate for individual supernovae, and always by $\lesssim 0.03$ dex in IMF-average—much smaller than errors from interpolation or the yield model itself. In any case, the specific estimate we use, or even individual abundance ratios, does not meaningfully impact our results.

As a test of consistency, we plot in Fig. 2 the IMF-averaged core-collapse supernova yields from NuGrid versus the pre-existing FIRE-2 yields (based on Nomoto et al. 2006). For almost all elements (except N) we find that IMF-averaging largely suppresses metallicity dependence; overall, a FIRE-2 default supernova produces 70 per cent more metals than one with NuGrid-based yields. Yields

\footnote{Core-collapse supernovae in 8–12 Myr old star particles (accounting for $\sim 7$ per cent of events), specifically at $Z = 10^{-4}$ or $10^{-3}$, show a difference between defined and estimated $[\alpha/Fe]$ of $\sim 0.3$ dex. This is still substantially smaller than the range in $[\alpha/Fe]$ of a star particle’s core-collapse ejecta over all time, which varies from 0.88 to 3.00 dex (0.91 to 3.34 dex in our estimate), with higher values associated with lower metallicities.}
Figure 1. Upper row shows aggregate properties for Type II core-collapse supernovae in NuGrid at the five metallicities it samples. Supernova rate is expressed in events per Myr per $M_\odot$ in a star particle, as a function of star particle age. Lower two rows show mass fractions of selected metals in core-collapse ejecta, as a function of time. For most progenitor masses, N exhibits a roughly linear metallicity dependence (as it does in the pre-existing FIRE-2 treatment, though this is not shown in the figure); in other cases, the dependence is usually moderate in strength and non-monotonic.

Figure 2. The IMF-averaged Type II core-collapse supernova yields of the FIRE-2 default and NuGrid-based yield prescriptions. While most yields are similar, production of O, Ne, and Mg differs substantially between the models. IMF-averaged metallicity dependence in NuGrid is weak, except for the production of N (which scales linearly with metallicity in default FIRE-2, although for clarity we omit this on the plot).

for H, He, C, Si, S, and Ca, agree between the two models to within ~20 per cent. For other elements the discrepancies are more pronounced: O production is reduced in NuGrid by a factor of ~2–3 depending on metallicity, while that of Ne is smaller by a factor of ~3–5. Mg yields are ~2–4 × lower, while those for Fe is larger by a factor of ~1–2. Consequently, in an IMF-averaged sense, the [$\alpha$/Fe] of core-collapse supernova ejecta is lower by ~0.3 dex in NuGrid with respect to FIRE-2-default models. While these differences in normalization impact overall trends in galactic elemental abundances, they do not change the increased dispersion in [$\alpha$/Fe] due to progenitor-mass dependence.

The MFM hydrodynamics method used here has no mass fluxes between fluid elements, so without additional modelling, metals deposited into a gas element would be locked there permanently, not mixing with surrounding gas elements. As a consequence, random variations in metal enrichment due to particle configuration, or which star particles have supernovae of a given mass when, would add to create substantial noise in a galaxy’s [$\alpha$/Fe] and [Fe/H] distributions. This is unphysical; realistically, marginally-unresolved fluid turbulence and instability should exchange metals between adjacent fluid elements and suppress this noise. Therefore, FIRE-2 simulations explicitly model subgrid metal diffusion/mixing (Hopkins 2017). Escala et al. (2018) showed that for MFM, metal diffusion is
required to eliminate noise and make detailed comparisons between simulated and observed [α/Fe] and [Fe/H] in galaxies; their results are robust with respect to the specific choice of diffusion coefficient. Williamson, Martel & Kawata (2016) obtain similar results for metal diffusion in smoothed-particle hydrodynamics (SPH) simulations, where there are also no fluxes between fluid elements. These properties are especially important for the simulations we present here, due to substantial variation in the [α/Fe] of core-collapse ejecta from different-mass progenitors1.

3 RESULTS

3.1 Star formation rates

We plot the SFRs as a function of lookback time for the three simulated galaxies in the upper panel of Fig. 3; in all cases, the SFR is bursty, peaking at $z \approx 4$ and declining thereafter. In each case, the cumulative SFH is available in Fig. 4. Both m10q.nugrid and m10q.nugrid_avg produce $\sim 10^6 M_\odot$ of stars over their lifetime, with m10q.nugrid_avg quenching its star formation at $z \approx 2$. m10q.FIRE-2, however, forms $\sim 2.7 \times 10^6 M_\odot$, with $\sim 30$ per cent of that mass being formed after $z = 2$. The differences in SFHs between runs are consistent with stochastic run-to-run variation that has been seen in other investigations of this particular galaxy (El-Badry et al. 2018; Hopkins et al. 2018; Su et al. 2018; Keller et al. 2019; Wheeler et al. 2019), due to the effects of random seeds and non-deterministic MPI reduction. We demonstrate this further in Fig. A1, where we show the cumulative fractional SFHs for all three runs. The yield model has no consistent effect on SFH or total stellar mass formed. We thus conclude that the choice between FIRE-2-default or NuGrid feedback parameters (supernova rate, energy, and ejecta mass), as well as including progenitor mass dependence, does not have a qualitatively significant impact on the dynamics and star formation of the simulated galaxies.

3.2 Metallicity distributions

In the lower panel of Fig. 3, we plot the $z = 0$ stellar-mass-weighted metallicity distribution function (MDF) for [Fe/H] in each of our simulations. All MDFs peak at an [Fe/H] $\approx -2.2$, approximately in line with the observed MDF of Ursa Minor, due to the inclusion of metal diffusion in the runs (Escala et al. 2018). However, the average metallicities (in log space; see Escala et al. 2018, for a definition) are lower, and standard deviations higher, in m10q.nugrid and m10q.nugrid_avg than in m10q.FIRE-2, with $<\text{[Fe/H]}>_{\text{NuGrid}} = -2.6 \pm 0.7$; these figures reflect both the lower level of star formation in the NuGrid runs, and the fatter low-metallicity tails of the NuGrid MDFs, including many stars at the metallicity ‘floor’ of the simulation ([Fe/H] $\sim -4$). Physically, the tails result from stochastic spatial and temporal variation in star formation—because metals are not yet well mixed by diffusion at early times, there is an increased scatter in [Fe/H] within and between the initial episodes of star formation, whose magnitude can differ from run to run.

To better understand metal enrichment, we also study the mean and dispersion in metallicity for individual starbursts (defined when star formation crosses $1 M_\odot$/Myr) in the middle panel of Fig. 3. We find that the typical iron abundance of each starburst, $<[\text{Fe/H}]>_s$, increases with time for all simulations, although it is lower by roughly 0.17 dex between m10q.nugrid and m10q.FIRE-2 due to lower star formation in the former. For starbursts at lookback times less than 10 Gyr (equivalently, $z < 2$), we find that $\sigma_{\text{[Fe/H]}_s} \approx 0.05$ dex for simulations with default FIRE-2 yields, and $\sim 0.1$–0.2 dex for

$^1$The m10q.nugrid_avg run does not show starbursts at late times with which to make a comparison, but early on follows the same pattern as the m10q.nugrid.
In the upper panel of Fig. 5, we show the full distribution (shaded regions), mean value (solid lines) and 1σ scatter (dotted lines) of [α/Fe] versus [Fe/H] for stars in all our simulations. We find, on average, that the mean abundance patterns for m10q.nugrid and m10q.nugrid_avg agree with one another, particularly at high metallicity. For both m10q.FIRE-2 and m10q.nugrid_avg, we find that the trend is very narrow, with a scatter of only 0.05 dex throughout the metallicity range we test for both stars and gas. With the fully time-dependent yields in m10q.nugrid, however, the scatter rises substantially to σ[α/Fe] ≈ 0.1–0.3 dex.3 This occurs because, with full-NuGrid yields, a star particle at a given [Fe/H] can produce supernovae with multiple possible [α/Fe] abundances, as opposed to just one with IMF-averaged yields.

Not only are non-IMF-averaged yields more conceptually realistic, they also better reflect the [α/Fe]-[Fe/H] features observed in some dwarf galaxies. While a detailed comparison of intrinsic scatter between theory and observation is not appropriate due to the widely varying methods for determining systematic error between samples, we can comment on the broad trends seen in some observations that would likely be better matched by mass- and metallicity-dependent yields. Aoki et al. (2009) observe a sample of six extremely metal-poor stars in Sextans, finding for the first time, a star with low [α/Fe] at low [Fe/H], hinting at a larger scatter in alpha abundances than is found in the halo of the Milky Way. This can also be seen in Sculptor and Carina (Kirby et al. 2009; Lemarque et al. 2012). Additionally, there is evidence that the M31 dwarf spheroidal galaxies may have more pronounced scatter than dSph satellites of the MW (Vargas et al. 2014; Kirby et al. 2020). Using a method similar to that outlined in Escala et al. (2018), we find the intrinsic scatter in [α/Fe] for the five dSph satellites presented in Kirby et al. (2020) to be 0.1–0.3 dex, while scatter in [α/Fe] for the MW satellites from Kirby et al. (2011a) is consistent with being entirely due to observational error. While we note that the error found for M31 satellites is similar to the scatter seen in m10q.nugrid, we caution that the method of estimating systematic error may add significant uncertainty to these results (Kirby, private communication).

m10q.nugrid also exhibits an increasingly pronounced intrinsic scatter for lower-metallicity stars. This trend is difficult to observe in dwarf galaxies, due to the low number of stars observed, but is seen in observations of halo stars (Ryan, Norris & Beers 1996), and is at least hinted at in some dwarf galaxy observations at low metallicity (Vargas et al. 2013, 2014; see also anomalous stars referenced in Frebel & Norris 2015). We note that m10q.nugrid broadly matches the observed trend of decreasing [α/Fe] with increasing [Fe/H], although flat trends have also been observed in some galaxies (Shetrone et al. 2001; Aoki et al. 2009; Frebel & Tafelmeyer et al. 2010).

Revaz et al. (2016) show that scatter in [α/Fe] versus [Fe/H] (specifically [Mg/Fe]) is artificially increased at sufficiently high resolution (m_{\text{bary}} \lesssim 1000 M_\odot) in simulations when stochastic sampling of the IMF and mass-dependent yields are used. They show that the scatter can be brought back down to observed levels (σ ≲ 0.3 dex for an ensemble of observed dwarf galaxies), by including either a metal mixing or a metal smoothing scheme. All of our runs include turbulent metal mixing, have high mass resolution, and all have σ ≲ 0.3. However, the lower panel of Fig. 5 shows a striking difference in the scatter in [α/Fe] at fixed [Fe/H] between m10q.FIRE-2 and m10q.nugrid_avg when compared to m10q.nugrid. This difference in scatter is also seen when using a second set of runs for each yield model (see Appendix A), demonstrating that variations in the IMF have a much larger impact on scatter in [α/Fe] versus [Fe/H] than does run-to-run variation. This is important, as run-to-run variation can be seen to drastically affect the SFH and stellar mass of a given run (see Fig. A1). This means that, for a given star born in a dwarf galaxy, when and under what galactic conditions it was born may have less of an effect on its abundances than the mass of stars that exploded in its neighbourhood prior to its birth.

In NuGrid, Type II core-collapse supernova ejecta from the most massive progenitors (with lifetimes ≲ 8 Myr, as seen in Fig. 1) exhibit a high ratio of [α/Fe] despite low overall production of α-elements and Fe. Ejecta from lower-mass progenitors, by contrast, contain substantially more Fe and therefore a lower [α/Fe] ratio. In m10q.nugrid, both types of supernovae enrich gas, and (through star formation that occurs more rapidly than turbulent diffusion eliminates [α/Fe] gradients) become represented on the [α/Fe] sequence. In m10q.nugrid_avg, however, IMF-averaging means that Fe-poor (high-mass) yields are pre-mixed with the Fe-rich (low-mass) yields, and so have little effect on the final [α/Fe] of the ejecta. Between −3.5 ≤ [Fe/H] ≤ −2.5, this effect causes [α/Fe] to be systematically ∼0.1 dex higher in m10q.nugrid than in m10q.nugrid_avg.

More substantial is the difference in α-abundances between the m10q.FIRE-2 and m10q.nugrid_avg simulations, with −[α/Fe]_{FIRE-2} ≈ −[α/Fe]_{NuGrid_avg} ≈ 0.2–0.4 dex, depending on [Fe/H]—in agreement with the expected IMF-averaged difference in [α/Fe] production per core-collapse supernova between Nomoto et al. (2006) and NuGrid (see Section 2). Observations suggest that for

3Note that the sharp spike in σ[α/Fe] for this run near [Fe/H] ∼ −1 is an artefact of the low-number of stars at that metallicity.
most stars in most dwarf galaxies, [$\alpha$/Fe] at low [Fe/H] is similar to
the abundance of the Galactic halo (Tolstoy et al. 2009; Frebel et al. 2010; Kirby et al. 2010; Hill et al. 2019). However, there
are some instances of stars having [$\alpha$/Fe] near or even below solar
at extremely low [Fe/H] (Shetrone et al. 2001; Aoki et al. 2009; Tafelmeyer et al. 2010; Starkenburg et al. 2013; Vargas et al. 2013),
more closely matching the abundances in m10q.nugrid. Of course
this could be high scatter masquerading as a lower normalization, but
the low number of stars at low [Fe/H] in dwarfs makes this difficult
to determine. We conclude that, despite capturing a larger dispersion
in [$\alpha$/Fe] versus [Fe/H], and matching some observed low values of
[$\alpha$/Fe] at low [Fe/H], the NuGrid tables may struggle to reproduce
the observed $\alpha$ abundance values at low metallicity found in most
dwarf galaxies.

4 CONCLUSIONS
We present the a detailed case study of cosmological hydrodynamic
simulation run with different SNe properties (energy, mass, and
yields) by comparing a single dwarf galaxy run with the default FIRE-
2 models (m10q.FIRE-2; Hopkins et al. 2018) to a more detailed
model that allows for complicated progenitor-mass and metallicity
dependence of all of the above from NuGrid (m10q.nugrid). We
also include a run in which we use the yields from NuGrid, but
average over the IMF (m10q.nugrid_avg).

We find that allowing for variations in the mass, metallicity, and
energy of individual explosions has only a weak effect on integral
galaxy properties (e.g. stellar masses, SFRs, total metallicities, MZR,
etc.), well within the range of stochastic variations. This is not
surprising, as these properties are primarily sensitive to other integral
quantities (e.g. total energy of SNe), as shown in the detailed study
of Su et al. (2018) who considered much more radical explosion-to-
explosion variations (e.g. hypernovae with $\sim 10^3 \times$ higher energy,
as opposed to factor $<2$ variations in SNe energy in the NuGrid
models).

As expected, the spread in detailed abundances or [$\alpha$/Fe] ratios
at a given time – or equivalently, value of [Fe/H] – is larger when
we include progenitor mass- and metallicity-dependent SNe yields.
Moreover, with the more detailed models, the scatter in [$\alpha$/Fe] increases
towards lower metallicity, in a manner similar to that
observed in some dwarf galaxies, as the enrichment becomes more
sensitive to the effect of individual SNe (as compared to a well-
mixed system which has been enriched by many SNe, and therefore
reflects an IMF-averaged population). For the systems here, above
[Fe/H] $\gtrsim -1$, IMF-averaging appears reasonable in the detailed

Figure 5. Above: probability densities in [$\alpha$/Fe] versus [Fe/H] space for star particles in our simulations (m10q.FIRE-2, blue; m10q.nugrid, black/grey; m10q.nugrid_avg, red). Solid lines indicate average [$\alpha$/Fe] at constant [Fe/H], while dotted lines indicate the $\pm 1\sigma$ intrinsic-scatter interval. Below: a more detailed plot of $1\sigma$ intrinsic scatter; the $\sigma \approx 0.1–0.3$ dex we find for m10q.nugrid is indicated in some observations (see the main text).
abundance tracks, while below this, IMF averaging likely produces too little variation in individual stellar abundance patterns. Thus, for future studies of metal-poor stellar populations, it will be important to include more detailed enrichment models of the sort studied here.

However, although the width and shape of the [$\alpha$/Fe] distribution and MDF in our detailed NuGrid models agrees reasonably well with observations, the absolute value of [$\alpha$/Fe] is systematically lower than that observed in most dwarf galaxies. This owes simply to the fact that even pure core-collapse SN yields (at any progenitor metallicity $Z < -1$), in IMF-average, give sub-solar [$\alpha$/Fe] in the NuGrid models. This is primarily driven by a combination of relatively high Fe yields for core-collapse explosions, and much lower (factor $\sim 3$–5) Mg, compared to other models such as those in Nomoto et al. (2006) as adopted in FIRE-2. If future observations of [$\alpha$/Fe] in extremely low metallicity stars continue to suggest a normalization $\sim 0.3$–0.6 dex higher than [m10q.nugrid] and [m10q.nugrid_avg], it is possible that some of NuGrid’s nucleosynthetic yields may need to be recalibrated.

ACKNOWLEDGEMENTS

We thank Benoit Côté, Andrew Graus, Alexander Ji, Evan Kirby, Mariska Kriek, Shea Garrison-Kimmel, Michael Grudic, Andrew Graus, Xiangcheng Ma, and Kung-Yi Su for useful discussions. Support for DM and coauthors was provided by NSF Collaborative Research Grants 1715847 and 1911233, NSF CAREER grant 1455342, and NASA grants 80NSSC18K0562 and JPL 1589072. Numerical calculations were run on the Caltech compute cluster ‘Wheeler,’ allocations FTA-Hopkins supported by the NSF and TACC, and NASA HEC SMD-16-7592. The data used in this work were, in part, hosted on facilities supported by the Scientific Computing Core at the Flatiron Institute, a division of the Simons Foundation. CRW acknowledges support from NASA through the NASA Hubble Fellowship grant #HST-HF2-51449.001-A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. AW received support from NASA through ATP grant 80NSSC18K1097 and HST grants GO-14734, AR-15057, AR-15809, and GO-15902 from STScI; a Scialog Award from the Heising-Simons Foundation; and a Hellman Foundation Fellowship. DK was supported by NSF Grant AST-1715101 and by a Cottrell Scholar Award from the Research Corporation for Science Advancement.

DATA AVAILABILITY

The data underlying this work will be shared at reasonable request to the corresponding author.

REFERENCES

Aoki W. et al., 2009, A&A, 502, 569
Cescutti G., 2008, A&A, 481, 691
Cohen J. G., Huang W., 2010, ApJ, 719, 931
Davé R., Thompson R., Hopkins P., 2016, MNRAS, 462, 3265
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
El-Badry K. et al., 2018, MNRAS, 473, 1930
Emeric A., Bryan G. L., Mac Low M.-M., 2019, MNRAS, 482, 1304
Escala I. et al., 2018, MNRAS, 474, 2194
Few C. G., Courty S., Gibson B. K., Kawata D., Calura F., Teyssier R., 2012, MNRAS, 424, L11
Frebel A., Norris J. E., 2015, ARA&A, 53, 631
Frebel A., Simon J. D., Geha M., Willman B., 2010, ApJ, 708, 560
Hassellquist S. et al., 2017, ApJ, 845, 162
Hill V. et al., 2019, A&A, 626, A15
Hirai Y., Saitoh T. R., 2017, ApJ, 838, L23
Ho N., Geha M., Tollerud E. J., Zinn R., Guhathakurta P., Vargas L. C., 2014, ApJ, 798, 77
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., 2017, MNRAS, 466, 3387
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Jeon M., Belsa G., Bromm V., 2017, ApJ, 848, 85
Kawata D., Arimoto N., Cen R., Gibson B. K., 2006, ApJ, 641, 785
Keller B. W., Wadsley J. W., Wang L., Kruijssen J. M. D., 2019, MNRAS, 482, 2244
Kirby E. N., Guhathakurta P., Bolte M., Sneden C., Geha M. C., 2009, ApJ, 705, 328
Kirby E. N. et al., 2010, ApJS, 191, 352
Kirby E. N., Lanfranchi G. A., Simon J. D., Cohen J. G., Guhathakurta P., 2011a, ApJ, 727, 78
Kirby E. N., Cohen J. G., Smith G. H., Majewski S. R., Sohn S. T., Guhathakurta P., 2011b, ApJ, 727, 79
Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, ApJ, 779, 102
Kirby E. N., Gilbert K. M., Escala I., Wojno J., Guhathakurta P., Majewski S. R., Beaton R. L., 2020, AJ, 159, 46
Kroupa P., 2002, Science, 295, 82
Lanfranchi G. A., Matteucci F., 2003, MNRAS, 345, 71
Leitherer C. et al., 1999, ApJS, 123, 3
Lemasle B. et al., 2012, A&A, 538, A100
Letarte B. et al., 2010, A&A, 523, A17
Mannucci F., Della Valle M., Panagia N., 2006, MNRAS, 370, 773
Maoz D., Graur O., 2017, ApJ, 848, 25
Marcolini A., D’Ercole A., Battaglia G., Gibson B. K., 2008, MNRAS, 386, 2173
Mosconi M. B., Tissera P. B., Lambas D. G., Cora S. A., 2001, MNRAS, 325, 34
Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, Nucl. Phys. A, 779, 424
Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, MNRAS, 406, 208
Pignatari M. et al., 2016, ApJS, 225, 24
Pillepich A. et al., 2018, MNRAS, 475, 648
Revaz Y., Jablonka P., 2018, MNRAS, 482, 2244
Revaz Y., Arnaudon A., Nichols M., Bonvin V., Jablonka P., 2016, A&A, 588, A21
Ritter C., Côté B., Herwig F., Navarro J. F., Fryer C. L., 2018a, ApJS, 237, 42
Ritter C., Herwig F., Jones S., Pignatari M., Fryer C., Hirschi R., 2018b, MNRAS, 480, 538
Ryan S. G., Norris J. E., Beers T. C., 1996, ApJ, 471, 254
Schaye J. et al., 2015, MNRAS, 446, 521
Schroyen J., De Rijcke S., Koleva M., Cloet-Osselaer A., Vandenbroucke B., 2013, MNRAS, 434, 888
Shetrone M. D., Côté P., Sargent W. L. W., 2001, ApJ, 548, 592
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Su K.-Y. et al., 2018, MNRAS, 480, 1666
Tafelmeyer M. et al., 2010, A&A, 524, A58
Tolstoy E., Hill V., Tosi M., 2009, A&A, 47, 371
Tsujimoto T., Nomoto K., Yoshii Y., Hashimoto M., Yangida S., Thielemann F. K., 1995, MNRAS, 277, 945
Valcke S., de Rijcke S., Dejonghe H., 2008, MNRAS, 389, 1111
Vargas L. C., Geha M., Kirby E. N., Simon J. D., 2013, ApJ, 767, 134
Vargas L. C., Geha M. C., Tollerud E. J., 2014, ApJ, 790, 73
Venn K. A. et al., 2012, ApJ, 751, 102

MNRAS 508, 508–515 (2021)
APPENDIX A: RUN-TO-RUN VARIATION

As can be seen in Fig. 3, the duration of star formation varies substantially between the m10q.FIRE-2, m10q.nugrid, and m10q.nugrid_avg runs. We argue in Section 3.1 that these differences are consistent with stochastic run-to-run variation between simulations, and are not related to the yield model used. In Fig. A1, we plot cumulative fractional SFHs for the original three runs (solid lines) alongside an additional run of each model with identical physical initial conditions (dotted lines). The stellar mass at $z = 0$ for each run is consistent to within a factor of $\sim 3$, with no clear relationship to the yield model employed.

In Fig. A2, we plot $[\text{Fe/H}]$ distribution functions for our fiducial and replication runs. There is no clear dependence on the metal-enrichment prescription, which strongly suggests that stochastic variations in final stellar mass and quenching time drive the differences between runs. Despite this, however, we show in Fig. A3 that the $[\alpha/\text{Fe}]$ versus $[\text{Fe/H}]$ tracks are nearly identical for each run of a given yield model. This strongly suggests that the conclusions presented here are robust against stochastic run-to-run variation.

Figure A1. Cumulative SFHs for the m10q.FIRE-2, m10q.nugrid, and m10q.nugrid_avg runs (solid) discussed in the text, as well as the corresponding replication runs (dashed). There is no clear association between SFH and the yield model used, and total stellar mass formed agrees to within a factor of $\sim 3$ between all runs.

Figure A2. $[\text{Fe/H}]$ distributions for our fiducial (solid) and replication (dashed) simulations. Differences between runs, particularly the low-metallicity tail, are driven by stochastic variation in SFH, and show little correlation with the enrichment prescription used. Above $[\text{Fe/H}] \approx -1$, there is little star formation, leading to substantial sampling error in the distribution.

Figure A3. $[\alpha/\text{Fe}]$ versus $[\text{Fe/H}]$ sequences for our fiducial (solid) and replication (dashed) simulations, with 1$\sigma$ confidence intervals shaded; above $[\text{Fe/H}] \approx -1$, sampling error means the sequences are poorly resolved. Despite substantial run-to-run variation in SFH, the sequences are well-converged, and demonstrate that time-dependent yields contribute to a greater intrinsic scatter.