Title

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

Authors

Muley, Dhruv A
Wheeler, Coral R
Hopkins, Philip F
et al.

Publication Date

2020-08-11

Peer reviewed

MNRAS 000, 1–7 (2020) Preprint 13 August 2020 Compiled using MNRAS LATEX style file v3.0

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

Dhruv A. Muley[1,2]*, Coral R. Wheeler[3,4,5]†, Philip F. Hopkins [3], Andrew Wetzel [6], Andrew Emerick [3,5], Dušan Kereš[4]

[1]*Department of Astronomy, University of California, Berkeley*
[2]*Department of Physics and Astronomy, University of Victoria*
[3]*California Institute of Technology, Pasadena*
[4]*University of California, San Diego*
[5]*Carnegie Observatories, Pasadena*
[6]*Department of Physics and Astronomy, University of California, Davis*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

In hydrodynamic simulations, prevailing subgrid chemical-evolution models often use a single, "IMF-averaged" supernova yield, ignoring variations in elemental abundance ratios (particularly [α/Fe]) in the ejecta of higher- and lower-mass supernova progenitors within a stellar population. To understand the impact of this simplification and understand the impact of more explicit models, we run FIRE simulations of a dwarf galaxy ($M_\star$(z = 0) $\sim 10^6\,M_\odot$) using nucleosynthetic yields from the `NuGrid` database 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 substantially widens the intrinsic scatter in the simulated [Fe/H]-[α/Fe]—a phenomenon potentially visible in recent 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 with ever-increasing precision (Kirby et al. 2009, 2010; Weisz et al. 2014; Ross et al. 2015). Abundance measurements from medium-resolution spectroscopy (e.g. Kirby et al. 2010) 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 eight Milky Way dwarf spheroidal (dSPh) satellite galaxies (Kirby et al. 2011a,b), and similar analyses have subsequently been performed on satellites of M31 (Vargas et al. 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, narrow metallicity distributions, and a generally decreasing trend in [α/Fe] (including Mg, Si, Ca, and Ti, per Kirby et al. 2011a) versus [Fe/H] (Kirby et al. 2013; Ho et al. 2014; Vargas et al. 2014). 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 (e.g., Ma et al. 2016; Escala et al. 2018; Wheeler et al. 2019) has improved our understanding of timescales and yields for processes such as AGB winds, Type Ia supernovae, and Type II core-collapse supernovae.

While simulations are broadly successful in matching the aggregate mass-metallicity relation of more massive ($M_* \gtrsim 10^9 M_\odot$) galaxies (Ma et al. 2016; Wetzel et al. 2016; Hopkins et al. 2018; Torrey et al. 2019), they have been less effective at reproducing more detailed abundance ratios (Sawala et al. 2010; Escala et al. 2018). In particular, they predict a very narrow sequence of [α/Fe] versus [Fe/H]; while this picture is broadly accurate, observations of some dwarfs (e.g., Vargas et al. 2014; Hasselquist et al. 2017; Kirby et al. 2020) may indicate a much greater intrinsic scatter, particularly at low metallicities/early times when Type II supernovae dominate enrichment (e.g., Tolstoy et al. 2009; Kirby et al. 2013; Frebel & Norris 2015). One potential reason for the discrepancy is that, in most theoretical work, Type II supernovae are assumed to have identical "IMF-averaged" yields, releasing mass, metallicity and energy in amounts independent of the age of the single stellar population repre-

*E-mail: dmuley@berkeley.edu
†Hubble Fellow

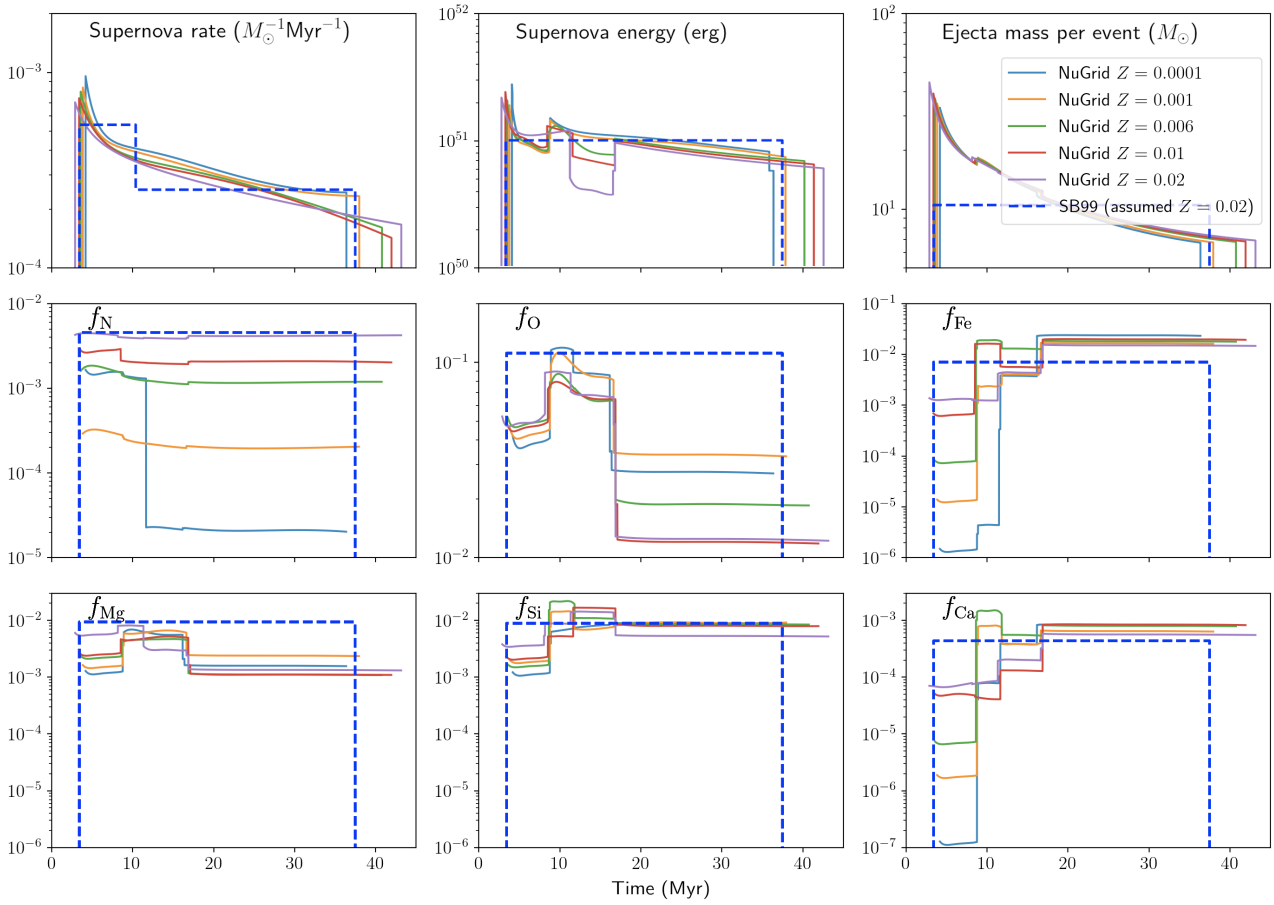

Figure 1. *Upper row*, aggregate properties for Type II core-collapse supernovae in `NuGrid` at the five metallicities it samples. Supernova rate is expressed in events per Myr per $M_\odot$ in a star particle, as a function of star particle age. *Lower two rows* show mass fractions of selected metals 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 FIRE2 treatment, though this is not shown in the figure); in other cases, the dependence is usually moderate in strength and non-monotonic.

sented by the star particle (although see e.g., Cescutti 2008; Wiersma et al. 2009, for exceptions). This unphysically obfuscates the inherent variation in abundance ratios between high-mass progenitors (which explode at early times) and lower-mass ones (which explode later).

In this paper, we present a case-study of the above effects 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 FIRE2 yields (**m10q.fire2**, based on Leitherer et al. 1999; Nomoto et al. 2006; Hopkins et al. 2018), the progenitor-mass- (and metallicity-) dependent `NuGrid` database 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 [α/Fe]-[Fe/H] to levels commensurate with 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 [α/Fe] abundance than observed, 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 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, timestep 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**,

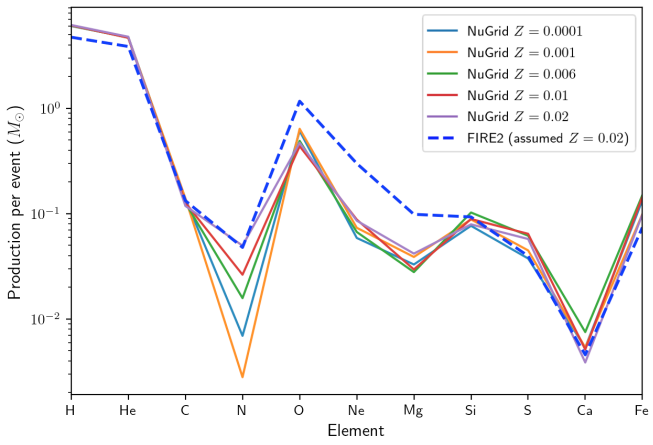

Figure 2. *Above*, the IMF-averaged Type II core-collapse supernova yields of the FIRE2 default and `NuGrid`-based yield prescriptions. While most yields are similar, production of O, Ne, and Mg differs substantially between the models. IMF-averaged metallicity dependence in `NuGrid` is weak, except for production of N (which scales linearly with metallicity in default FIRE2, although for clarity we omit this on the plot).

we average these yields over the IMF. On average, `NuGrid` core-collapse supernovae produce $\sim 13 M_\odot$ of ejecta per event, while the default FIRE2 core-collapse supernovae produce 10.5 $M_\odot$. Given the timesteps 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 $R_{\rm Type\ Ia} \propto t^{-1.1}$ delay-time distribution (from Maoz & Graur 2017) rather than the Gaussian prompt+constant delay treatment (Mannucci et al. 2006) used in **m10q.fire2** and in previous FIRE2 simulations. This increases the number of Type Ias by $\sim$2×, to $\sim$15% of all supernovae.

In the top row of Figure 1, we plot the rates, energies, and ejecta mass for core-collapse supernova events in `NuGrid`, alongside the `STARBURST99`-based (Leitherer et al. 1999) FIRE2 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%, 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 star-formation rates 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 FIRE2 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 ($f_{\rm N} \approx 4.05\times10^{-3}(Z/Z_\odot)$), with the exception of massive progenitors at the lowest metallicity $Z = 10^{-4}$.

Of the four α-elements comprising the Kirby et al. (2011a) observational definition, we directly track Mg, Si, and Ca, plotted in the bottom row of Figure 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 [(Mg, Si, Ca)/Fe] of 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 $[\alpha/{\rm Fe}] \equiv (1/4)([{\rm Mg/Fe}] + [{\rm Si/Fe}] + [{\rm Ca/Fe}] + [{\rm Ti/Fe}]) \approx (1/4)([{\rm Mg/Fe}] + [{\rm Si/Fe}] + 2[{\rm Ca/Fe}])$). We motivate this choice by the fact that in the `NuGrid` tables, $[\alpha/{\rm Fe}]$ for core-collapse ejecta almost[1] 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 (e.g., $[\alpha/{\rm Fe}] \approx (1/3)([{\rm Mg/Fe}] + [{\rm Si/Fe}] + [{\rm Ca/Fe}])$, or even individual abundance ratios) does not meaningfully impact our results.

As a test of consistency, we plot in Figure 2 the IMF-averaged core-collapse supernova yields from `NuGrid` versus the pre-existing FIRE2 yields (based on Nomoto et al. 2006). For almost all elements (except N) we find that IMF-averaging largely suppresses metallicity dependence; overall, a FIRE2 default supernova produces 70% more metals than one with `NuGrid`-based yields. Yields for H, He, C, Si, S, and Ca, agree between the two models to within $\sim$20%. For other elements the discrepancies are more pronounced: O production is reduced in `NuGrid` by a factor of $\sim 2-3$ depending on metallicity, while that of Ne is smaller by a factor of $\sim 3-5$. Mg yields are $\sim 2-4\times$ lower, while those for Fe is larger by a factor of $\sim 1-2$. Consequently, in an IMF-averaged sense, the $[\alpha/{\rm Fe}]$ of core-collapse supernova ejecta is lower by $\sim$0.3 dex in `NuGrid` with respect to FIRE2-default models. While these differences in normalization impact overall trends in galactic elemental abundances, they do not change the increased *dispersion* in $[\alpha/{\rm Fe}]$ due to progenitor-mass dependence.

The MFM hydrodynamics method used here has no mass fluxes between fluid elements, so without additional modeling, metals deposited into a gas element would remain there permanently and not mix with surrounding gas elements. As a consequence, random variations in metal enrichment due to particle configuration, and in the particles

[1] Core-collapse supernovae in 8-12 Myr old star particles (accounting for $\sim$7% of events), specifically at $Z = 10^{-4}$ or 10^{-2}, show a difference between defined and estimated $[\alpha/{\rm Fe}]$ of $\sim$0.3 dex. This is still substantially smaller than the range in $[\alpha/{\rm Fe}]$ 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.

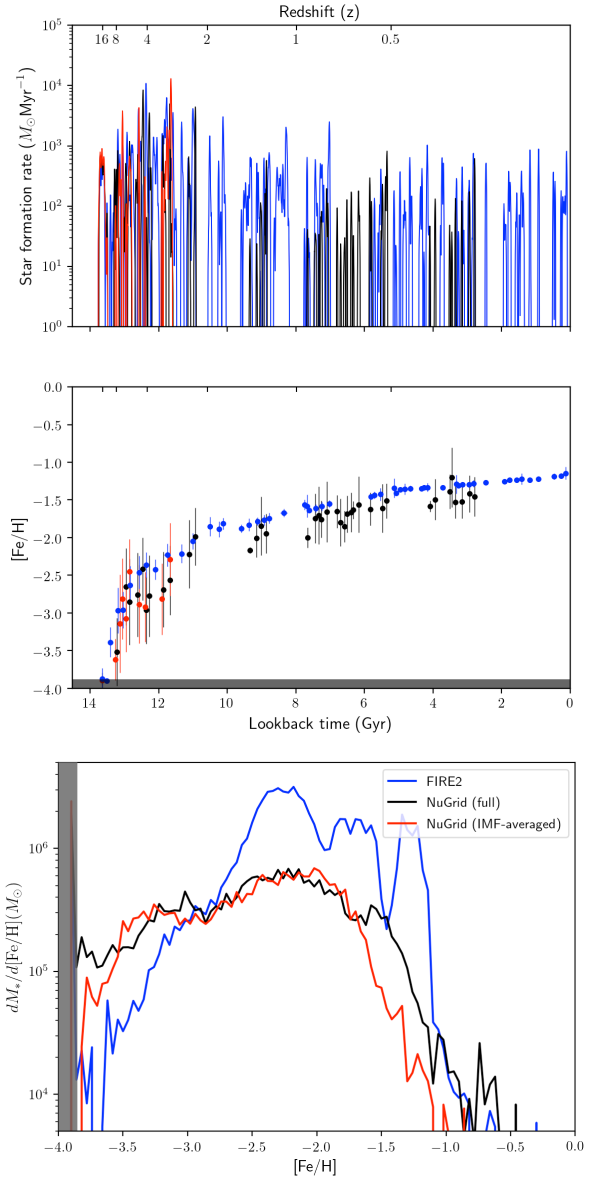

Figure 3. *Above*, the star-formation rate in each of our simulations over time, smoothed on a 20 Myr timescale; the `NuGrid` runs stochastically self-quench by $z \sim 2$, as in default-FIRE2 runs (e.g., Su et al. 2018). *Middle*, the mean (points) and standard deviation (lines) in [Fe/H] for stars formed during each starburst, with the `NuGrid` runs showing substantially greater intrinsic scatter; in all cases, however, this decreases with time as turbulent diffusion mixes metals. *Below*, non-normalized [Fe/H] MDF of our simulations. In the lower two panels, grey areas indicate the region of the "metallicity floor" ([Fe/H]~ -4) in our simulations, where the earliest-formed stars artificially accumulate.

that form stars, would stack to create substantial noise in a galaxy's [α/Fe] and [Fe/H] distributions. This is unphysical; realistically, marginally-unresolved processes such as turbulence and instability should exchange metals between adjacent fluid elements and suppress this noise. Therefore, FIRE-2 simulations explicitly model sub-grid metal diffusion/mixing (Hopkins 2017). Without this, Escala et al. (2018) showed that detailed comparisons between simulated and observed [α/Fe] and [Fe/H] become intractable—doubly true in our case due to substantial variation in the [α/Fe] of core-collapse ejecta from different-mass progenitors [1].

3 RESULTS

3.1 Star-formation rates

We plot the star-formation rates (SFRs) as a function of lookback time for the three simulated galaxies in the upper panel of Figure 3; in all cases, the SFR is bursty, peaking at $z \sim 4$ and declining thereafter. 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.fire2**, however, forms $\sim 2.7 \times 10^6 M_\odot$, with $\sim$30% of that mass being formed after $z = 2$. The differences in star formation histories 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 Appendix A. We thus conclude that the choice between FIRE2-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 Figure 3, we plot the stellar-mass-weighted metallicity distribution function (MDF) for [Fe/H] in each of our simulations. All MDFs peak at an [Fe/H] ≈ -2.2. 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.fire2**, with $\langle$[Fe/H]$\rangle_{\rm FIRE2} = -2.1 \pm 0.5$, while $\langle$[Fe/H]$\rangle_{\rm NuGrid\ full} = -2.5 \pm 0.7$ and $\langle$[Fe/H]$\rangle_{\rm NuGrid\ avg} = -2.6 \pm 0.7$; these figures reflect the fatter low-metallicity tails of the `NuGrid` MDFs, including many stars at the metallicity "floor" of the simulation ([Fe/H]~ -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 Figure 3. We find that the typical iron abundance of each starburst, $\langle$[Fe/H]$\rangle_{\rm sb}$, increases with time for all simulations, although it is lower by roughly 0.17 dex between **m10q.nugrid** and **m10q.fire2** due to lower star formation in the former.[2] For starbursts at lookback times less than 10 Gyr (equivalently, $z \lesssim 2$), we find that $\sigma_{\rm [Fe/H],sb} \approx 0.05$ dex for simulations with default FIRE2

[2] 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**.

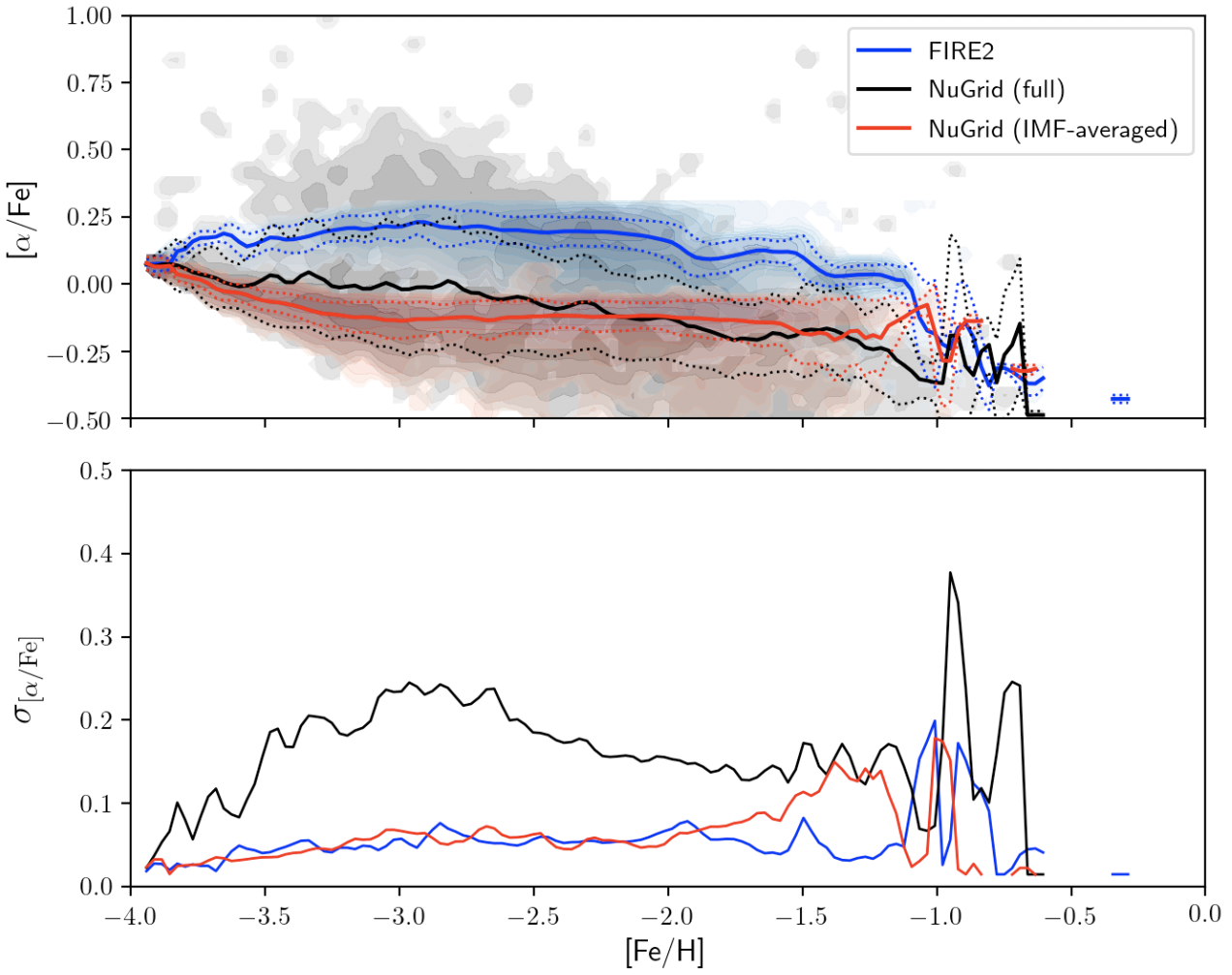

Figure 4. *Above*, probability densities in [α/Fe] vs. [Fe/H]-space for star particles in our simulations (**m10q.fire2** *blue*, **m10q.nugrid** *black/grey*, **m10q.nugrid_avg** *red*). Solid lines indicate average [α/Fe] at constant [Fe/H], while dotted lines indicate the $\pm1\sigma$ intrinsic-scatter interval. *Below*, a more detailed plot of 1σ intrinsic scatter; the $\sigma \approx 0.1-0.3$ dex we find for **m10q.nugrid** is indicated in some observations (see main text).

yields, and $\sim 0.1-0.2$ dex for the full-`NuGrid` simulation, in line with observations; this intrinsic scatter decreases with time in all simulations (Kirby et al. 2013; Escala et al. 2018) as turbulent diffusion brings metal concentrations into equilibrium. There is a positive correlation between $\sigma_{\rm [Fe/H],sb}$ and burst duration for **m10q.nugrid**, which is absent in the other (IMF-averaged) runs; this suggests that short-term (e.g., time-dependent SNe II metal yields), long-term (e.g., Type Ia and AGB enrichment), and numerical factors (turbulent diffusion, stochastic star formation) all contribute to the measured intrinsic scatter. To leading order, however, the width and shape of the MDF is predominated by differences in metallicity *between* stellar populations formed at different times.

3.3 Alpha-element abundances

In the upper panel of Figure 4, we show the full distribution (shaded regions), mean value (solid lines) and $1-\sigma$ 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.fire2** and **m10q.nugrid_avg**, we find that the trend is very narrow, with a scatter of only $\sigma_{\rm [\alpha/Fe]} \approx 0.05$ 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 $\sigma_{\rm [\alpha/Fe]} \approx 0.1-0.3$ dex. 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. Hasselquist et al. (2017)'s high-resolution APOGEE spectroscopy of the Sagittarius dSph demonstrate high intrinsic scatter, albeit at higher [Fe/H] than that seen here due to its $M_* \approx 4\times10^8\,{\rm M}_\odot$. Within the mass range we study, Hill et al. (2019)'s VLT/FLAMES observations of [Ca/Fe] in the Sculptor dSph, as well as medium-resolution Keck studies of M31 satellites (Vargas et al. 2014; Kirby et al. 2020), indi-

cate an intrinsic scatter of 0.1-0.3 dex (although for M31, the method of estimating systematic error may add uncertainty to these results: Kirby 2020, private communication). We also reproduce the more pronounced intrinsic scatter for lower-metallicity stars (e.g., Tolstoy et al. 2009; Frebel & Norris 2015), as expected for stars formed before a galaxy's metals have mixed evenly. We note that **m10q.nugrid** broadly matches the observed trend of decreasing [α/Fe] with increasing [Fe/H].

In `NuGrid`, Type II core-collapse supernova ejecta from the most massive progenitors (with lifetimes $\lesssim$8 Myr, as seen in Figure 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 $-3.5 \lesssim$ [Fe/H] $\lesssim -2.5$, 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.fire2** and **m10q.nugrid_avg** simulations, with $\langle[\alpha/\mathrm{Fe}]\rangle_{\mathrm{FIRE2}} - \langle[\alpha/\mathrm{Fe}]\rangle_{\mathrm{NuGrid\ avg}} \approx 0.2 - 0.4$ 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 Sec. 2). We thus conclude that, despite capturing a larger *dispersion* in [α/Fe] versus [Fe/H], the `NuGrid` tables cannot accurately reproduce the normalization found in observations (e.g., Kirby et al. 2011a).

4 CONCLUSIONS

We present the a detailed case-study of cosmological hydrodynamic simulations run with different SNe properties (energy, mass, and yields) by comparing a single dwarf galaxy run with the default FIRE-2 models (**m10q.fire2**; Hopkins et al. 2018) to a more detailed model which 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 weak effects 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 $\sim 1000x$ higher energy, as opposed to factor < 2 variations in SNe energy in the `NuGrid` models).

As expected, the spread in detailed abundances or [α/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 [α/Fe] *increases* towards lower metallicities, in a manner similar to that observed, as the enrichment becomes more sensitive to the effective 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 abundance tracks, while below this, IMF averaging produces far 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 [α/Fe] distribution and MDF in our detailed `NuGrid` models agrees reasonably well with observations, the absolute value of [α/Fe] is systematically lower than that observed by $\sim$ 0.3–0.6 dex. This owes simply to the fact that even *pure* core-collapse SNe yields (at *any* progenitor metallicity $Z < -1$), in IMF-average, give sub-solar [α/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. This implies 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 Grudić, Andrew Graus, Xiangcheng Ma, and Kung-Yi Su for useful discussions. Support for DM and coauthors was provided by NSF Collaborative Research Grants 1715847 & 1911233, NSF CAREER grant 1455342, NASA grants 80NSSC18K0562, JPL 1589742. 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; the Heising-Simons Foundation; and a Hellman Fellowship. DK was supported by NSF Grant AST-1715101 and by a Cottrell Scholar Award from the Research Corporation for Science Advancement.

REFERENCES

Cescutti G., 2008, A&A, 481, 691
El-Badry K., et al., 2018, MNRAS, 473, 1930
Escala I., et al., 2018, MNRAS, 474, 2194
Frebel A., Norris J. E., 2015, ARA&A, 53, 631
Hasselquist S., et al., 2017, ApJ, 845, 162
Hill V., et al., 2019, A&A, 626, A15

Ho N., Geha M., Tollerud E. J., Zinn R., Guhathakurta P., Vargas L. C., 2014, The Astrophysical Journal, 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
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
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
Mannucci F., Della Valle M., Panagia N., 2006, MNRAS, 370, 773
Maoz D., Graur O., 2017, ApJ, 848, 25
Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, Nuclear Phys. A, 777, 424
Pignatari M., et al., 2016, ApJS, 225, 24
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
Ross T. L., Holtzman J., Saha A., Anthony-Twarog B. J., 2015, AJ, 149, 198
Sawala T., Scannapieco C., Maio U., White S., 2010, MNRAS, 402, 1599
Su K.-Y., et al., 2018, MNRAS, 480, 1666
Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371
Torrey P., et al., 2019, MNRAS, 484, 5587
Vargas L. C., Geha M. C., Tollerud E. J., 2014, ApJ, 790, 73
Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, ApJ, 789, 147
Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23
Wheeler C., et al., 2019, MNRAS, 490, 4447
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Wojno J., Gilbert K. M., Kirby E. N., Escala I., Beaton R. L., Tollerud E. J., Majewski S. R., Guhathakurta P., 2020, ApJ, 895, 78

APPENDIX A: RUN-TO-RUN VARIATION

As can be seen in Figure 3, the duration of star formation varies substantially between the **m10q.fire2**, **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 Figure A1, we plot cumulative fractional star formation histories (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 ~3, with no clear relationship to the yield model employed. Despite the stochasticity in final stellar mass and quenching time (from currently star forming to quenched by $z \sim 3$), we show in Figure A2 that the [α/Fe] versus [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.

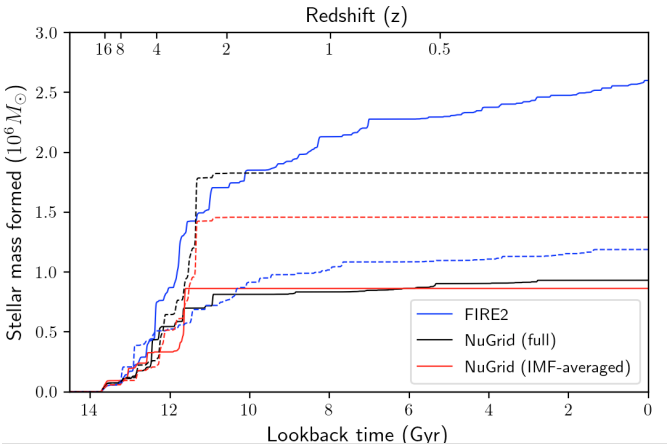

Figure A1. Cumulative fractional star-formation histories for the **m10q.fire2**, **m10q.nugrid**, and **m10q.nugrid_avg** runs *(solid)* discussed in the text, as well as the corresponding replication runs *(dashed)*. There is no clear association between SFH and the yield model used, and total stellar mass formed agrees to within a factor of ~3 between all runs.

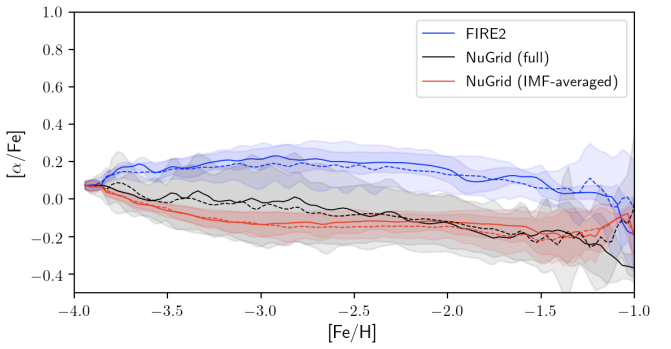

Figure A2. [α/Fe] versus [Fe/H] sequences for our fiducial *(solid)* and replication *(dashed)* simulations, with 1-sigma confidence intervals shaded; above [Fe/H] ≈ -1, sampling error means the sequences are poorly resolved. Despite substantial run-to-run variation in star-formation history, the sequences are well-converged, and demonstrate that time-dependent yields contribute to a greater intrinsic scatter.

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