

The challenge of simultaneously matching the observed diversity of chemical abundance patterns in cosmological hydrodynamical simulations

Tobias Buck¹,¹★ Jan Rybizki², Sven Buder^{3,4}, Aura Obreja⁵, Andrea V. Macciò^{2,6,7},
Christoph Pfrommer¹, Matthias Steinmetz¹ and Melissa Ness⁸

¹Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

²Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

³Research School of Astronomy and Astrophysics, Australian National University, ACT 2611, Australia

⁴Center of Excellence for Astrophysics in Three Dimensions (ASTRO-3D), Australia

⁵Universitäts-Sternwarte München, Scheinerstrasse 1, D-81679 München, Germany

⁶New York University Abu Dhabi, Saadiyat Island, PO Box 129188 Abu Dhabi, United Arab Emirates

⁷Center for Astro, Particle and Planetary Physics (CAP3), New York University, PO Box 129188 Abu Dhabi, United Arab Emirates

⁸Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY, USA

Accepted 2021 September 20. Received 2021 September 20; in original form 2021 March 1

ABSTRACT

With the advent of large spectroscopic surveys the amount of high quality chemodynamical data in the Milky Way (MW) increased tremendously. Accurately and correctly capturing and explaining the detailed features in the high-quality observational data is notoriously difficult for state-of-the-art numerical models. In order to keep up with the quantity and quality of observational data sets, improved prescriptions for galactic chemical evolution need to be incorporated into the simulations. Here we present a new, flexible, time-resolved chemical enrichment model for cosmological simulations. Our model allows us to easily change a number of stellar physics parameters such as the shape of the initial mass function (IMF), stellar lifetimes, chemical yields, or SN Ia delay times. We implement our model into the *Gasoline2* code and perform a series of cosmological simulations varying a number of key parameters, foremost evaluating different stellar yield sets for massive stars from the literature. We find that total metallicity, total iron abundance, and gas phase oxygen abundance are robust predictions from different yield sets and in agreement with observational relations. On the other hand, individual element abundances, especially alpha-elements show significant differences across different yield sets and none of our models can simultaneously match constraints on the dwarf and MW mass scale. This offers a unique way of observationally constraining model parameters. For MW mass galaxies we find for most yield tables tested in this work a bimodality in the $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ plane of rather low intrinsic scatter potentially in tension with the observed abundance scatter.

Key words: methods: numerical – galaxy: abundances – galaxy: structure – galaxies: abundances – galaxies: evolution – galaxies: formation.

1 INTRODUCTION

Galaxies are the lighthouses in the vast ocean of the Universe (Mo, van den Bosch & White 2010). Their structure and appearance is the result of the interplay of several generations of stars which are shaped by different epochs of gas accretion and expulsion. Therefore galactic structures encode valuable information about the fundamental physical processes governing galaxy formation. Especially the abundance of chemical elements in stars, the interstellar medium (ISM), and the circumgalactic medium (CGM) serves as an ideal tracer of the integrated galactic history.

The observed diversity of chemical elements with atomic numbers larger or equal to C is the result of stellar nucleosynthesis and the distribution of those elements in the Universe a consequence of gas expulsion from star forming regions and galaxies via processes dubbed feedback (e.g. Maiolino & Mannucci 2019). Those chemical

elements are ideal observational tracers to determine the evolutionary state of stars or the physical conditions such as density, temperature, or ionization state of the gas. They further contribute importantly to the dynamics of the whole system as main coolants which enable the dissipation of energy by the emission of radiation. Thus the production and evolution of heavy elements is one of the fundamental ingredients of modern models for the formation and evolution of both stars and galaxies (e.g. Gibson et al. 1997; Chiappini et al. 2003; Mollá & Díaz 2005; Martínez-Serrano et al. 2008; Vogelsberger et al. 2020).

The stellar lifetimes and the production of new chemical elements (in the reminder of this article referred to as metals) strongly depends on the mass of a star and its own chemical composition. When a generation of stars evolves it will pollute the ISM with chemical elements fused in their cores, predominantly during their last evolutionary phases. As a consequence, the successive generations of stars born from the enriched ISM will attain a higher mass fraction of heavy elements with respect to their progenitors. Thereby chemical evolution does not proceed in a linear manner due to the strong mass dependence of stellar evolution time-scales. For example, high mass stars

* E-mail: tbuck@aip.de

evolve much more quickly compared to low and intermediate mass stars ($\tau \sim 4$ Myr for a $40 M_{\odot}$ star and $\tau \sim 10$ Gyr for a $1 M_{\odot}$ star).

Therefore, higher mass stars first enrich the ISM with heavy chemical elements while intermediate and low mass stars lock the chemical elements produced in past generations for much longer times (up to 1–10 Gyr, e.g. Marigo 2001; Thielemann et al. 2003) with the immediate consequence that the chemical characteristics of second-generation stars will be mainly influenced by the dying massive stars. Consequently, including the (time resolved) formation, distribution, exchange, and recycling of chemical elements into any simulation of the formation and evolution of galaxies is critical for our understanding.

Since the first implementation of chemical enrichment by the metal release of core-collapse supernovae (CC-SN) into a smoothed particle hydrodynamics (SPH) code (Steinmetz & Mueller 1994) much progress on the implementation of chemical evolution within cosmological simulations has been made (e.g. Raiteri, Villata & Navarro 1996; Kawata 2001; Mosconi et al. 2001; Recchi, Matteucci & D’Ercole 2001; Lia, Portinari & Carraro 2002; Tornatore et al. 2007; Oppenheimer & Davé 2008; Hirschmann, De Lucia & Fontanot 2016). By now there exist several models that distinguish between elements originating from SN Ia and CC-SN and in some cases also from asymptotic-giant branch (AGB) stars and the winds from their progenitors (e.g. Scannapieco et al. 2005; Brooks et al. 2007; Tissera et al. 2012; Brook et al. 2014; Naiman et al. 2018) taking into account the self-consistent calculation of the metallicity depend gas cooling rates (Martínez-Serrano et al. 2008; Wiersma et al. 2009). Recently, Valentini et al. (2019) studied different model assumptions including IMF variations, stellar yields, and binarity on a simulated MW-like galaxy, and found that these imprint important differences in the resulting object.

Additionally, much progress in the collection of high quality observational data has been made and a correlation between the observed abundance patterns of chemical elements in stars and the ISM with the mass of a galaxy, the position within a galaxy and the birth epoch of the stars (e.g. Baade 1944; Wallerstein 1962; Hayden et al. 2015) could be established. Furthermore, there exist a great wealth of data for the chemical abundances of the CGM at low- and high-redshift. But most notably with the advent of large spectroscopic surveys the amount of high quality chemodynamical data in the Milky Way (MW) increased tremendously (Boeche et al. 2011; Anders et al. 2014; Buder et al. 2019; Xiang et al. 2019; Steinmetz et al. 2020) exhibiting distinct chemical abundance patterns of the stellar disc (e.g. Boeche et al. 2013; Hayden et al. 2015) and even enabling us to reconstruct MW’s spiral pattern (e.g. Eilers et al. 2020). Because the surface abundance of stars is assumed to reflect the chemical composition of the star’s birth gas cloud the measured chemical abundances present a fossil record of the Galaxy’s formation history. The study of the chemokinematic structures of the MW in combination with inferred stellar ages (Bensby et al. 2014; Feuillet et al. 2016; Martig et al. 2016; Mackereth et al. 2017; Ness et al. 2019) enabled the field of Galactic Archeology (Freeman & Bland-Hawthorn 2002) and has lead us to unique insights into the formation and evolution of the MW’s stellar disc (e.g. Minchev, Chiappini & Martig 2013; Sestito et al. 2021).

Accurately capturing and explaining the detailed features present in the observational data is notoriously difficult, but recent numerical models begin to reproduce observed features (e.g. Doménech-Moral et al. 2012; Miranda et al. 2016; Domínguez-Tenreiro et al. 2017; Grand et al. 2018; Hilmi et al. 2020; Khoperskov et al. 2020; Agertz et al. 2021; Renaud et al. 2021a). By construction chemical evolution models encompass a large number of stellar physics parameters such

as the shape of the initial mass function (IMF), stellar lifetimes or chemical yields, or SN Ia delay times. Many of those parameters are poorly constrained a priori (Côté et al. 2016b) and chosen ad-hoc. Recently, Gutcke & Springel (2019) for instance studied the influence of a metallicity-dependent high mass slope of the IMF on galactic abundance patterns. Perhaps the most important external or theoretical input for chemical evolution models are the nucleosynthetic yields for the various enrichment channels (Romano et al. 2010; Côté et al. 2016a). In the past, theoretical yields have produced miss-matches with the observations and the physical shortcomings of stellar nucleosynthetic yield models are still under debate (e.g. Nomoto, Kobayashi & Tominaga 2013; Müller 2016; Pignatari et al. 2016).

Thus, in order to cope with the large range of parameters to explore, various galactic chemical evolution (GCE) models have been developed (e.g. Tinsley 1979; Chiappini, Matteucci & Gratton 1997; Gibson 1997; Rybizki, Just & Rix 2017; Ritter et al. 2018a; Kobayashi, Karakas & Lugaro 2020). These models range from analytical ones with simplifying assumptions (e.g. Spitoni, Vincenzo & Matteucci 2017; Weinberg, Andrews & Freudenburg 2017; Spitoni et al. 2020) to hydrodynamical simulations including detailed chemical enrichment as well as galactic dynamics (e.g. Kawata & Gibson 2003; Few et al. 2012; Gibson et al. 2013; Brook et al. 2020; Buck et al. 2020). All these numerical models made possible to test a large set of parameters over a wide range of values, and marginalize over parameters for simple astrophysical inferences. Yet, due to their simplifying assumptions those models lack the possibility of including all the complex dynamics present in full hydrodynamical simulations which might potentially explain some of their shortcomings as seen Blancato et al. (e.g. in 2019).

Therefore, in this work we present a new hybrid approach of chemical evolution models and fully cosmological simulations of galaxy formation where essential chemical evolution parameters are first constrained by fitting a galactic chemical evolution model to observational data (as e.g. done in Philcox, Rybizki & Gutcke 2018). The obtained values for essential chemical evolution parameters are then employed in the chemical enrichment prescription in cosmological hydrodynamical simulations of the formation of galaxies from dwarf to MW mass. The scientific aim of this work is twofold: (i) we update the chemical enrichment prescription of the *Gasoline2* code to be capable of tracing many different chemical elements and (ii) benchmark the results obtained by assuming different yield tables from the literature. In this way, we constrain the systematic uncertainties on galactic abundance patterns arising from differences in the stellar evolution models.

This paper is organized as follows: in Section 2, we present the details of the chemical evolution model followed by the description of a novel implementation of chemical enrichment into a cosmological simulation code in Section 3. Readers which are already familiar with the concept of simple stellar population models and their usage in cosmological simulations might readily skip those two sections. We continue in Section 4 where we analyse the results of the numerical simulations with special emphasize on the abundance patterns of the stellar discs in light of current galactic archeological studies. We end this study with a discussion and summary of our results in Section 5.

2 METHOD: A FLEXIBLE CHEMICAL EVOLUTION MODEL

State-of-the-art computational resources still do not enable cosmological simulations to resolve individual stars. Instead, the standard approach is to discretize the stellar density field into tracer particles

which represent a population of stars. The population of stars contained in those stellar particles is described by a single age, τ_{star} , a metallicity, Z_{star} , and an IMF specifying the number of stars in a given mass bin. Such stellar populations are usually referred to as simple stellar population (SSP). In order to determine how the SSP interacts with the surrounding ISM one has to determine the (time resolved) amount of energy, metals, and mass released by the stellar particle, usually referred to as stellar feedback.

There are many stellar evolution models in the literature employing slightly different stellar lifetimes functions, IMFs, or metal mass return fractions. Instead of deciding for one specific model we make use of the flexible chemical evolution code Chempy¹ (Rybizki et al. 2017) in order to synthesize the final stellar evolution model used in the simulations. This means we are using Chempy to fix the SSP parameters to synthesize the yield tables to be used for a specific run. In future work, Chempy can be further used to fit the SSP parameters to observed data and thus constrain the SSP models employed in hydrodynamical simulations.²

In summary, Chempy is a general purpose tool to link the parameters of a stellar (chemical) evolution model θ (e.g. the IMF high-mass slope and/or a specific yield set) to the likelihood of an observations \mathcal{O} (such as stellar abundances) via its model predictions d . A detailed prescription of the modelling procedure is given in the code paper (Rybizki et al. 2017) but for completeness we describe the basic working principle of Chempy below. Since our aim is to use Chempy to create an SSP model which we then can utilize as an input to cosmological simulations we focus here on the SSP part of Chempy while only briefly explaining its ISM parameters which are of little interest for us.

2.1 The simple stellar population model

At the core Chempy calculates the time evolution of the chemical yields for an SSP. In order to specify the stellar physics of the chemical evolution model at least seven parameters need to be set (see Table A1) and additional parameters for the stellar nucleosynthetic yields have to be chosen (see Table 1). Currently those include the relative mass return fractions from three nucleosynthetic channels: Supernova of type Ia (SN Ia), core-collapse supernova (CC-SNe), and asymptotic giant branch stars (AGB).

An SSP is fully characterized by its age, mass, and initial element composition, $\text{SSP}(\tau_{\text{star}}, m, [\text{X}/\text{H}])$. By assuming some IMF, a stellar lifetime function and specific nucleosynthetic yields, the stellar (energetic and mass) feedback is simultaneously fixed. In the case of numerical simulations the total mass of an SSP is set by the stellar particle mass and its initial elemental abundance is inherited from the composition of the ISM from which the star particle is formed (see also Saitoh 2017, for a compilation of models). For this work, we are focusing on fixed feedback fractions and do not consider stochastically sampling the IMF, though it can be accounted for in Chempy.

The simple stellar population model consists of five ingredients: (i) the stellar initial mass function, (ii) the stellar lifetime function, (iii) the (empirical) supernova Ia delay time distribution, (iv) the stellar mass-loss from stellar evolution models, and (v) the chemical

Table 1. Yield tables implemented in CHEMPY.

Yield table	Masses	Metallicities
CC SN		
Portinari, Chiosi & Bressan (1998)	[6,120]	[0.0004,0.05]
François et al. (2004)	[11,40]	[0.02]
Chieffi & Limongi (2004)	[13,35]	[0,0.02]
Nomoto et al. (2013)	[13,40]	[0.001,0.05]
Frischknecht et al. (2016)	[15,40]	[0.00001,0.0134]
West & Heger (in prep.)	[13,30]	[0,0.3]
Ritter et al. (2018b)	[12,25]	[0.0001,0.02]
Limongi & Chieffi (2018) ³	[13,120]	[0.0000134,0.0134]
SN _{Ia}		
Iwamoto et al. (1999)	[1.38]	[0,0.02]
Thielemann et al. (2003)	[1.374]	[0.02]
Seitenzahl et al. (2013)	[1.40]	[0.02]
AGB		
Karakas (2010)	[1,6.5]	[0.0001,0.02]
Ventura et al. (2013)	[1,6.5]	[0.0001,0.02]
Pignatari et al. (2016)	[1,65.5]	[0.01,0.02]
Karakas & Lugaro (2016)	[1,8]	[0.001,0.03]
TNG ⁴	[1,7.5]	[0.0001,0.02]
Hypernova		
Nomoto et al. (2013)	[20,40]	[0.001,0.05]

composition of that mass-loss. We detail our model assumptions and parameters for (i)–(iv) in Appendix A as those are standard SSP parameters and only discuss the most important part for this work, the stellar yields, here below.

2.1.1 Nucleosynthetic yields

The elemental enrichment in Chempy is calculated according to publicly available yield tables from the literature. A list of implemented yield tables is presented in Table 1 but it is straightforward to extend this list further using Chempy's python functions. Our fiducial yield set for example combines tables by Karakas & Lugaro (2016) for AGB stars with tables from Chieffi & Limongi (2004) for CC-SN and yields from Seitenzahl et al. (2013) for SN Ia⁵ (all combinations of yield tables used in this work are reported in Table 2).

Stellar nucleosynthesis changes the relative fraction of elements inside a star by different destruction and production rates for various elements. Light elements such as hydrogen or helium are preferentially destroyed in favour of more massive elements. Similarly massive elements such as iron do not take part in further fusion processes and can be regarded as the end product of stellar nucleosynthesis. Therefore we split the elemental enrichment of stars into two parts: the first part describes the amount of newly synthesized elements and the second part describes the initial abundance of elements which passes unaltered through the star. This second part describes the elemental composition of the birth ISM which is locked up in the star's atmosphere and released back into the ISM at the end of the

³Using the rotation parametrization of Prantzos et al. (2018).

⁴The Ilustris-TNG (Naiman et al. 2018; Pillepich et al. 2018) yield set for AGB stars is a mixture of yields taken from Karakas (2010), Doherty et al. (2014), and Fishlock et al. (2014).

⁵We use their N100 model calculated in 3D superseding the 1D W7 model of Iwamoto et al. (1999) which had old electron capture rates overproducing Ni and underproducing Mn abundances. This is remedied with the Seitenzahl et al. (2013) models which best reproduce observables by Sim et al. (2013).

¹A Python implementation of the current version of Chempy can be found on GitHub <https://github.com/jan-rybizki/Chempy>.

²Note, we use the part of Chempy that calculates the IMF integrated metal yield of an SSP. This is equivalent to SYGMA (Ritter et al. 2018a) in the NUPYCEE collaboration.

Table 2. Combinations of yield sets for CC-SN, SN Ia, and AGB star yields and SSP parameters probed in this work. Our fiducial model is referred to as *new fid* while other combinations varying CC-SN yields are referred to as *alt*, *alt2*, *alt3*. A variation of AGB star yields is referred to as *alt4*. Other physics parameter variations of the fiducial yield set are referred to as *steepIMF* for a steeper high mass slope of the IMF, *longDelay* for a longer SN Ia delay time and *highNorm* for a higher SN Ia normalization.

Name	CC-SN yield	SN Ia yield	AGB yield	IMF slope	SN Ia delay time	SN Ia normalization
new fid	Chieffi & Limongi (2004)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.3	40 Myr	−2.9
alt	Limongi & Chieffi (2018)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.3	40 Myr	−2.9
alt2	West & Heger (in prep.)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.3	40 Myr	−2.9
alt3	Nomoto et al. (2013)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.3	40 Myr	−2.9
alt4	Nomoto et al. (2013)	Seitenzahl et al. (2013)	TNG	−2.3	40 Myr	−2.9
steepIMF	Chieffi & Limongi (2004)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.7	40 Myr	−2.9
longDelay	Chieffi & Limongi (2004)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.3	158 Myr	−2.9
highNorm	Chieffi & Limongi (2004)	Seitenzahl et al. (2013)	Karakas & Lugaro (2016)	−2.3	40 Myr	−2.75

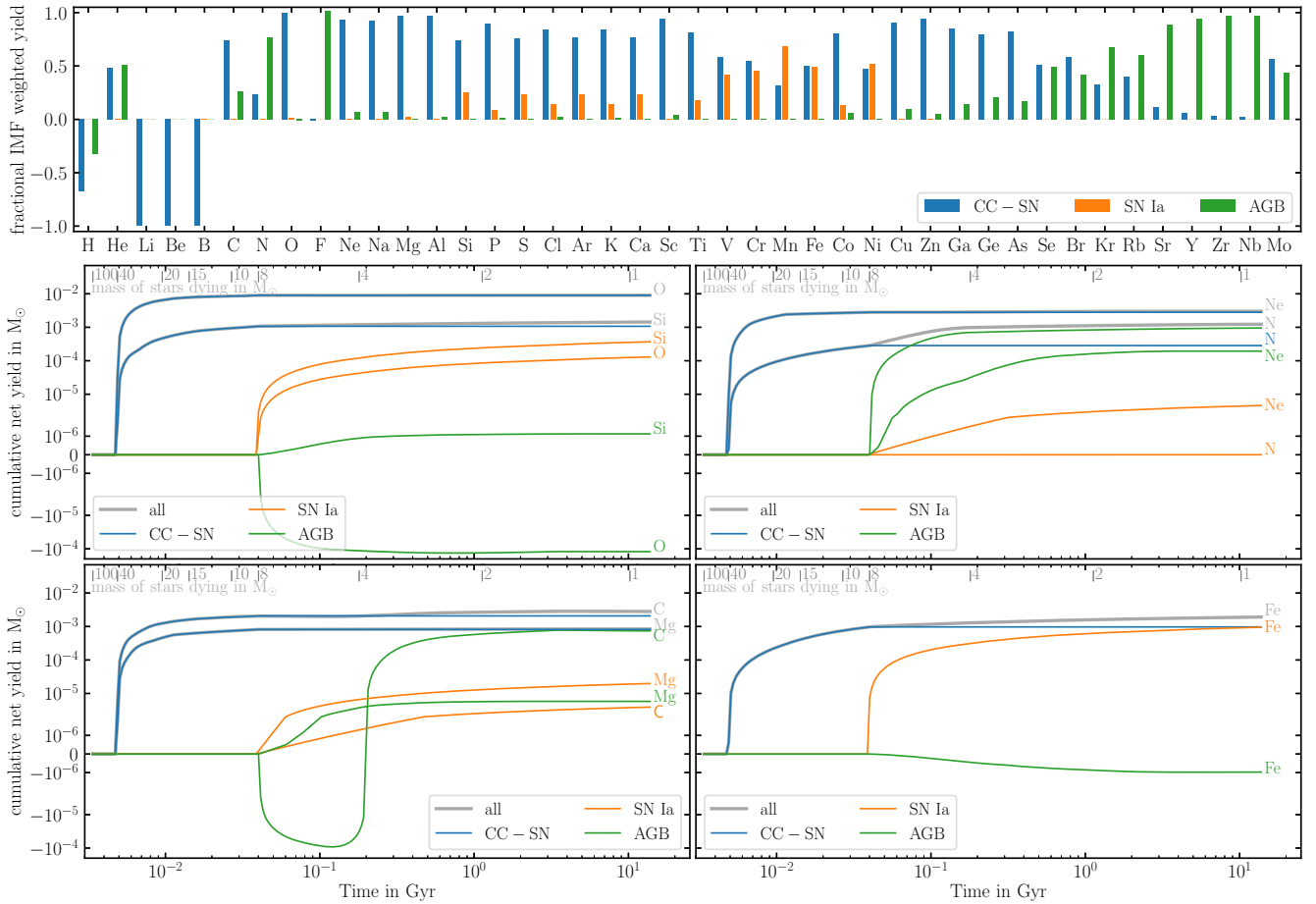


Figure 1. Example tracks of traced elements. The upper panel shows the fractional element return of the whole SSP for 42 example elements (out of a total of 81 tracked elements) split into the three different channels of elements produced by CC-SN (blue), AGB (green) stars, and SNIa (orange) after a Hubble time. The histograms show the relative contribution of different processes to the total metal return of the SSP; e.g. CC-SN and SN Ia contribute equally to the production of iron. Negative values as e.g. for hydrogen indicate destruction/fusion of elements. The four bottom panels show the element return of the SSP as a function of time for different elements as indicated in each panel. We split the total element return (grey line) into the three different channels showing the contribution by CC-SN in blue, by AGB stars in green, and by SNIa in orange. Note, the negative iron yield for AGB stars reflects the iron destruction through neutron capture process.

star's life (see also discussion in section 4 of Wiersma et al. 2009). We refer to the yield sets following this distinction of newly produced elements from the birth of elemental abundance of the stars as net-yields while the tables including the birth composition are referred to as gross-yields. Formulating the elemental mass return in this

way ensures that the yield tables are independent of the exact initial chemical composition of the underlying CC-SN models and contain only the newly synthesized elements. On the other hand, in the gross-yield formulation tables might contain elements which were neither synthesized in the star nor available in the ISM it was formed from

simply because they were part of the initial abundances of the CC-SN explosion model. All yield tables implemented in Chempy list the element return in the net-yield formulation and exist for various initial stellar metallicities (see Table 1).

Fig. A3 shows the cumulative time evolution of the net-yields of seven different elements returned by an SSP of total mass $1M_{\odot}$. Thick blue line show an SSP of solar metallicity while thin orange lines show results for an SSP of $10^{-5}Z_{\odot}$. These panels show that O, Ne, and Mg are all released early on by massive stars with short lifetimes. C, N, or Fe on the other hand have significant late time contributions from lower mass stars and SN Ia explosions. Comparing blue and orange lines, we can further appreciate significant evolution of the net-yields with stellar metallicity. Note, C is the only element which shows larger yields at low metallicity possibly explaining the observations of carbon-enhanced metal-poor stars in the MW (e.g. Starkenburg et al. 2014).

Fig. 1 shows in more detail the origin of the differences in the time release of elements for our fiducial yield set. In this figure, we split the cumulative yield return into the contributions from different nucleosynthetic channels (CC-SN in blue, SN Ia in orange, and AGB winds in green). The top panel shows a histogram of the fractional net-yield contribution of each nucleosynthetic channel integrated over the IMF and over a Hubble time of a subset of 42 out of a total of 81 elements traced by Chempy.⁶

The values for H, Li, Be, and B are negative indicating their preferential consumption inside stars. For the other elements, these figures show that most of them originate from a combination of different channels with varying relative contributions. For example, for iron and Ni we see that CC-SN and SN Ia contribute equal amounts of Fe to the enrichment (similar conclusions can be made for Cr, V). O, Na, Mg, Al, and Cu are almost only produced by CC-SN while F stems entirely from AGB winds. For He on the other hand we see that CC-SN and AGB winds contribute roughly equal amounts of mass to the enrichment. C is mostly produced in CC-SN but has roughly one quarter contribution from AGB winds while N shows the exact opposite. Mn again is mostly released by SN Ia with an one-third contribution by CC-SN. Finally, Si, S, Ar, or Ca originate from CC-SN explosions with roughly one quarter contribution by SN Ia.

While for some elements different channels might contribute equally to the elemental return there can be a significant time delay in the onset of the release of elements of up to 100 Myr (roughly one dynamical time inside the stellar disc of an L^* galaxy) between the release of elements from different sources. This is highlighted in the lower four panels showing the cumulative net-yields of different elements as a function of time (grey lines) split into the contributions from different nucleosynthetic channels (CC-SN in blue, SN Ia in orange, and AGB winds in green). While CC-SN start releasing their elements as soon as the most massive stars explode ($>40M_{\odot}$, $t_{\min} \sim 4$ Myr after birth of the SSP), SN Ia, and AGB star winds release their elements much later (~ 40 Myr after birth). Interestingly, the release of O by SN Ia explosions is roughly offset by the destruction of oxygen in AGB stars. Similarly there exists a period of time when the AGB net-yield of C is negative due to high mass AGB stars⁷ and only becoming positive when lower mass stars enter the AGB phase.

While the time evolution is qualitatively similar for all net-yields taken from different sources in the literature their quantitative agreement is not. There is no real consensus on how much mass of a given element is synthesized in a specific star (as a function of metallicity and mass) nor the relative contribution of the different channels to a single element. In fact, there can be a difference of up to a factor of a few between different yield sets. This fact is highlighted in Fig. 2 which compares the fractional difference between IMF averaged net-yields from different yield tables in the literature in comparison to our fiducial yield set. The upper panel compares SN Ia yields, the middle panel CC-SN yields (see also Fig. D1 for additional yield tables), and the lower panel compares AGB wind yields. Note, the y-axis is in logarithmic scale which shows that for some elements the difference can be up to a factor of a few.

At first sight this might give the impression that in general chemical yields might be unreliable. However, investigating this figure more closely we see that for the most abundant elements (H, He, O, C, Ne, Fe, N, Si, Mg, S) the differences between different yield tables is smaller than $\lesssim 20$ per cent across all channels (see also Romano et al. 2010, for a more in depth discussion of uncertainties related to specific stellar yields). This intrinsic uncertainty should be kept in mind when comparing model predictions to observations. We further note that those are only the uncertainties attributed to the stellar models and nucleosynthetic processes. Additional uncertainties originate when using a different IMF (see e.g. discussion of Fig. A1).

2.1.2 Coupling the SSP model to numerical simulations

In Figs A3 and 1, we have shown that there is substantial time evolution of the chemical enrichment for a single SSP over its lifetime. Therefore it is necessary to explicitly follow the time release of elements in numerical simulations instead of applying an instantaneous recycling approximation. The various elements follow distinct tracks which make it relatively complicated to accurately approximate their evolution by a small set of analytic functions. Therefore a lookup table approach seems the most straightforward implementation in order to make use of the chemical enrichment in numerical simulations.

The usage of net-yield tables in Chempy makes the application of the chemical enrichment prescriptions formulated in Chempy in numerical simulations relatively easy. Those simulations usually trace the initial elemental abundance of gas and stellar particles such that the combined elemental return for an element i , Δm_{tot}^i in a time-step Δt from a stellar particle of mass m_{star} follows via

$$\Delta m_{\text{tot}}^i = \Delta m_{\text{new}}^i + \Delta m_{\text{init}}^i, \quad (1)$$

where Δm_{new}^i is the newly synthesized mass of element i taken from the net-yield tables at the appropriate stellar metallicity that is released during this time-step and is the amount of initial mass of that element inherited from the birth gas cloud that is released in this time-step. The sum of Δm_{tot}^i over all elements i will sum up exactly to the total mass lost by the SSP. Note that Δm_{new}^i might be negative for elements such as hydrogen or helium which get fused into higher mass elements. Therefore, mass conservation demands that the sum over Δm_{new}^i of all elements i in the net-yield table should return 0 since the mass of newly synthesized heavy elements equals the mass of the destroyed lighter elements.

into higher mass elements effectively lowering the C abundance of the initial composition.

⁶See Tutorial 8 of the Chempy package at <https://github.com/jan-rybizki/Cchempy/tree/master/tutorials>

⁷Note, this is only true for the net-yields. The total mass return to the ISM will be the sum of newly synthesized elements referred to as the net-yields and the initial elemental abundance. Here, some of the initial C will be fused

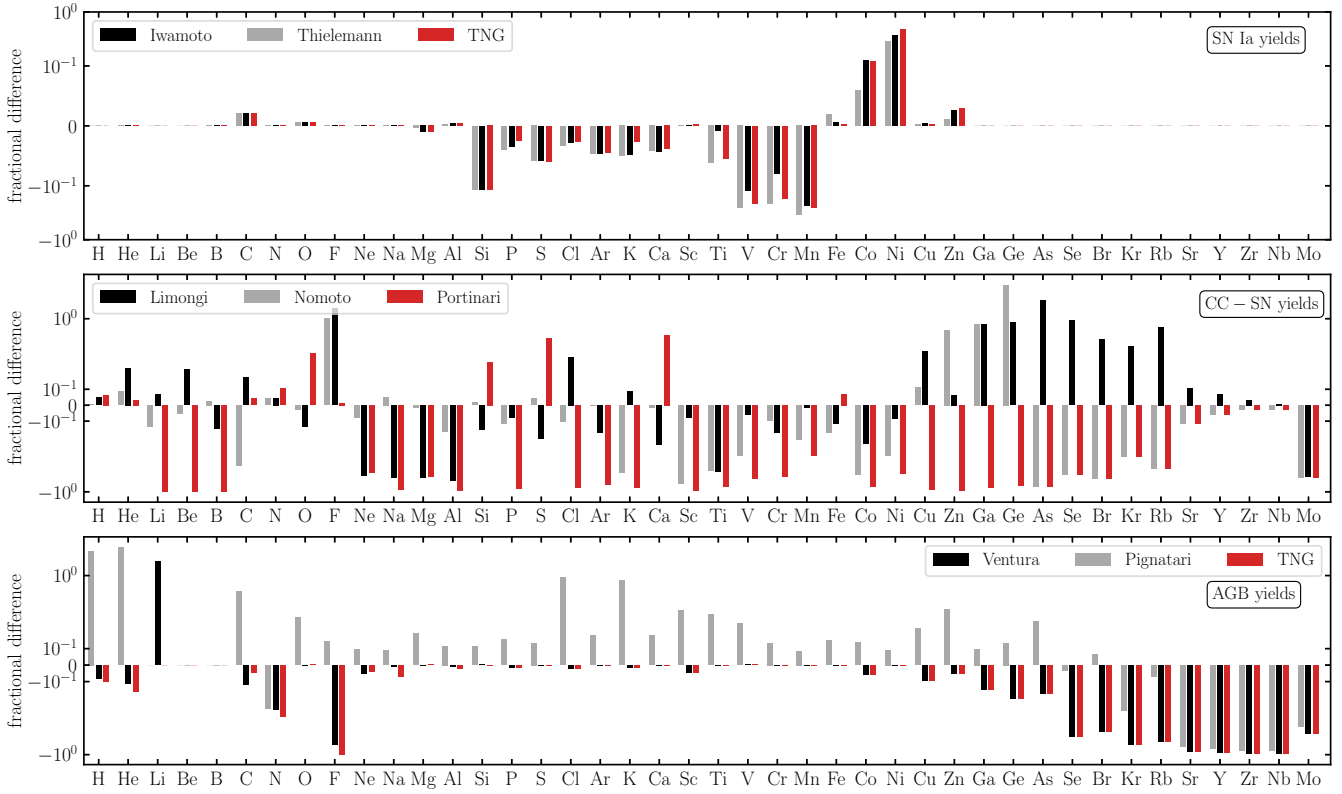


Figure 2. IMF weighted relative differences between our fiducial yield set using CC SN yields from Chieffi & Limongi (2004), AGB star yields from Karakas & Lugaro (2016), and SNIa yields from Seitzzahl et al. (2013) and other yield sets implemented in Chempy as indicated in the legend. The top panel shows SNIa yields by Iwamoto et al. (1999), Thielemann et al. (2003), and the Illustris-TNG yields (see the text for more explanation, (see text for more explanation, Naiman et al. 2018; Pillepich et al. 2018). The two middle panels compare CC SN yields from Limongi & Chieffi (2018), Nomoto et al. (2013), Portinari et al. (1998). The bottom panel compares AGB yields by Ventura et al. (2013), Pignatari et al. (2016), and the Illustris-TNG yields with our fiducial yield set. An additional comparison of CC SN yields from Frischknecht et al. (2016), West & Heger (in preparation) and the Nugrid collaboration (Ritter et al. 2018b) is shown in Fig. D1 in the appendix.

3 APPLICATION: COSMOLOGICAL ZOOM-IN SIMULATIONS

We have taken the above approach and implemented the time release of elements as synthesized by Chempy into the modern smoothed particle hydrodynamics (SPH) code *Gasoline2* (Wadsley, Keller & Quinn 2017). For this study, we perform in total 23 cosmological zoom-in simulations of 8 different galaxies ranging from 10^6 – $10^{11} M_{\odot}$ in stellar mass. Energetic feedback prescriptions and initial conditions are adopted from the NIHAO project (Wang et al. 2015) and for completeness we briefly discuss them below. Basic parameters of each central galaxy are listed in Table 3. Note, the projected stellar half-light radii measured for the galaxies are well in agreement with observed galaxy sizes at the same stellar mass for which we expect half-light sizes of ~ 3 kpc (e.g. Kravtsov 2013). For the simulated Milky Way analogues those correspond to exponential scale length of ~ 4 – 5 kpc (table 2 of Buck et al. 2020).

The NIHAO galaxies adopt cosmological parameters from the Planck Collaboration VIII (2014), namely: $\Omega_m = 0.3175$, $\Omega_{\Lambda} = 0.6825$, $\Omega_b = 0.049$, $H_0 = 67.1 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\sigma_8 = 0.8344$. Zoom-in initial conditions are created using a modified version of the GRAFIC2 package (Bertschinger 2001; Penzo et al. 2014) by selecting galaxies for re-simulation from cosmological dark matter only boxes selecting only on the dark halo mass in combination with an isolation criterion for numerical reasons. Formally, we only consider haloes that have no other halo with mass greater than one-fifth of the virial mass within three virial radii.

The adopted force softening for dark matter, gas, and star particles are listed in Table 3. Simulated galaxies from the NIHAO project generally match the observed properties of central galaxies (see e.g. Obreja et al. 2016, 2019a; Buck et al. 2017, 2018; Dutton et al. 2017) and MW/M31 satellites (Buck et al. 2019a). They further show good numerical convergence in terms of resolution (Buck et al. 2019c, 2020) and feedback prescriptions (e.g. Obreja et al. 2019b). In particular, the adopted feedback prescriptions result in the formation of a chemical bimodality of the stellar disc of L^* galaxies (Buck 2020).

3.1 Hydrodynamics

The hydrodynamics solver *Gasoline2* (Wadsley et al. 2017) employs a modern formulation of the SPH method which improves multiphase mixing and removes spurious numerical surface tension. The simulations employ the Wendland C2 smoothing kernel (Dehnen & Aly 2012) with a number of 50 neighbour particles for the calculation of the SPH properties. The treatment of artificial viscosity uses the signal velocity as described in Price (2008) and a time-step limiter as described in Saitoh & Makino (2009) has been implemented such that cool particles behave correctly when hit by a hot blastwave. All simulations employ a pressure floor following Agertz, Teyssier & Moore (2009) to keep the Jeans mass of the gas resolved by at least four SPH kernel masses to suppress artificial fragmentation. This is equivalent to the criteria proposed in Richings & Schaye (2016).

and fulfils the Truelove et al. (1997) criterion at all times. Finally, we adopted a turbulent metal diffusion algorithm between SPH particles as described in Wadsley, Veeravalli & Couchman (2008) and extensively discussed in Section 2.2 of Shen, Wadsley & Stinson (2010). This model estimates a particle's diffusion coefficient as $d = C_{\text{turb}} |S_{ij}| h^2$, where S_{ij} denotes the trace-free local velocity shear tensor defined as

$$\tilde{S}_{ij}|_p = \frac{1}{\rho_p} \sum_q m_q (v_{j|q} - v_{j|p}) \nabla_{p,i} W_{pq},$$

$$S_{ij}|_p = \frac{1}{2} (\tilde{S}_{ij}|_p + \tilde{S}_{ji}|_p) - \delta_{ij} \frac{1}{3} \text{Trace } \tilde{S}|_p. \quad (2)$$

The sums extend over all SPH neighbours, q , δ_{ij} denotes the Kronecker delta, ρ_q the density, \mathbf{r}_{pq} the vector separation between particles p and q , W is the SPH kernel function, $v_{i|q}$ is the particle velocity component in direction i and ∇_p is the gradient operator for particle p (operating on the SPH kernel function) while $\nabla_{p,i}$ is the i th component of the resultant vector. This choice for S_{ij} results in no diffusion for compressive or purely rotating flows. With equation (2) the diffusion expression for a scalar variable A_p of particle p can be computed as follows:

$$D_p = C |S_{ij}|_p h_p^2,$$

$$\frac{dA_p}{dt} = - \sum_q m_q \frac{(D_p + D_q)(A_p - A_q)(\mathbf{r}_{pq} \cdot \nabla_p W_{pq})}{\frac{1}{2}(\rho_p + \rho_q) \mathbf{r}_{pq}^2}. \quad (3)$$

Finally, the coefficient C_{turb} is set to 0.05 (Wadsley et al. 2008; Shen et al. 2010).

Gas cooling in the temperature range from 10 to 10^9 K is implemented via hydrogen, helium, and various metal-lines using pre-calculated `cloudy` (version 07.02; Ferland et al. 1998) tables (Shen et al. 2010; Obreja et al. 2019b). Heating from cosmic reionization is implemented via a metagalactic UV background (Haardt & Madau 2012).

3.2 Star formation and feedback

Star formation is implemented following Stinson et al. (2006) in which dense ($n_{\text{th}} > 50 m_{\text{gas}} / \epsilon_{\text{gas}}^3 = 10.3 \text{ cm}^{-3}$, where m_{gas} denotes the gas particle mass, ϵ_{gas} the gas gravitational softening, and the value of 50 refers to the number of neighbouring particles) and cold ($T < 15\,000$ K) gas is eligible to form stars. A systematic study of the impact of the star formation threshold density on the galaxy properties has recently been conducted by Dutton et al. (2019, 2020) where only simulations with a high ($n > 10 \text{ cm}^{-3}$) threshold density reproduce the observed spatial clustering of young star clusters (Buck, Dutton & Macciò 2019b).

The gas fulfilling the above requirements will be stochastically converted into star particles according to: $\frac{\Delta m_{\text{star}}}{\Delta t} = c_{\text{star}} \frac{m_{\text{gas}}}{t_{\text{dyn}}}$, where Δm_{star} is the mass of the star particle formed, m_{gas} the gas particle mass, Δt the time-step between star formation events (here: 8×10^5 yr), and t_{dyn} is the gas particles dynamical time. The variable c_{star} refers to the star formation efficiency, i.e. the fraction of gas that will be converted into stars during the time t_{dyn} . We set this parameter to $c_{\text{star}} = 0.1$ which reproduces the Kennicutt (1998) Schmidt relation (Stinson et al. 2006). The initial stellar particle mass at birth is fixed to one third of the initial gas mass and successive mass-loss is implemented as described in Section A4.

We model the energy input from stellar winds and photoionization from luminous young stars following the ‘early stellar feedback’ prescriptions in Stinson et al. (2013) and inject the resulting feedback energy as pure thermal feedback. The energy input from CC-

SN and SN Ia blastwaves is implemented following Stinson et al. (2006) using self-consistent supernovae rates calculated by the stellar evolution models described in Section 2. For both feedback channels, the efficiency parameters were chosen such that one MW mass galaxy respects the abundance matching relations (Behroozi, Wechsler & Conroy 2013; Kravtsov, Vikhlinin & Meshcheryakov 2018; Moster, Naab & White 2018) at all redshifts.

Elemental feedback from CC-SN, SN Ia, and AGB stars is implemented as described in Section 2.1.2. We tabulate the time-resolved, mass-dependent element release of a single stellar population as a function of initial metallicity in a grid of 50 metallicity bins logarithmically spaced between 10^{-5} and 0.05 in metallicity with each metal bin resolved by 100 time bins logarithmically spaced in time from 0 to 13.8 Gyr. All our simulations track the evolution of the 10 most abundant elements by default (H, He, O, C, Ne, Fe, N, Si, Mg, S) while any other element present in the yield tables can additionally be tracked. Our fiducial combination of yield tables uses SN Ia yields from Seitenzahl et al. (2013), CC-SN yields from Chieffi & Limongi (2004) and AGB star yields from Karakas & Lugaro (2016). Additionally we simulated a dwarf galaxy and a MW-mass galaxy with varying CC-SN yields (see Table 2 for more details) as those yields should have the most effect on the final chemical composition (see e.g. Fig. 1). The python code exemplifying how to use `Chempy` to synthesize the yield tables can be found at <https://github.com/TobiBu/chemical.enrichment.git>.

For this work we define the abundances of element X as $[X/H] = \log(m_X/m_H) - \log(X_{X,\odot}/X_{H,\odot})$ where m_X and m_H are the gas or star particle's mass of element X and hydrogen mass while $X_{X,\odot}$ and $X_{H,\odot}$ denote the solar mass fractions of element X and hydrogen taken from Asplund et al. (2009). Finally, the abundance of α elements is defined as the (mass weighted) sum of Mg, Si, Ca and Ti such that $[\alpha/\text{Fe}]$ is $[(\text{Mg}+\text{Si}+\text{Ca}+\text{Ti})/\text{Fe}]$. We compare our simulations to observational data from Kirby et al. (2019) and Buder et al. (2021).

4 RESULTS: GALACTIC CHEMICAL COMPOSITION

We evaluate the new chemical enrichment model implemented in `Gasoline2` by investigating the stellar mass–metallicity relation (left-hand panel of Fig. 3), the stellar $[\text{Fe}/\text{H}]$ versus stellar mass relation (right-hand panel of Fig. 3) and the gas phase oxygen abundance (Fig. 4). We compare our simulations to observations from Gallazzi et al. (2005), Panter et al. (2008), Kirby et al. (2013) and Tremonti et al. (2004), Zahid et al. (2013), and Berg et al. (2012), respectively. We further investigate changes with respect to the fiducial NIHAO suite.

4.1 Stellar mass–metallicity relation

The left-hand panel of Fig. 3 shows the stellar mass–metallicity relation for the new chemical enrichment model (coloured symbols) in comparison to observational data as indicated in the legend as well as results from the fiducial NIHAO implementation (faint blue dots). The overall shape of this relation is mainly set by the interplay of star formation, feedback, and gas inflow/outflow (e.g. Dayal et al. 2013). Stellar masses and metallicity are measured for all stars within two times the projected stellar half-light radius, $2 R_{\text{half}}$. The right-hand side panel shows the stellar iron abundance versus stellar mass both measured within the projected half-light radius of the galaxies in comparison with the relation for Local group dwarf galaxies measured by Kirby et al. (2013).

Table 3. Simulation properties of the main galaxies: We state the total stellar mass, M_{star} and the total amount of gas, M_{gas} , within the virial radius. We further report the projected stellar half-light radius, R_{half} , the total metallicity, Z/Z_{\odot} , iron abundance $[\text{Fe}/\text{H}]$, and gas phase oxygen abundance measured within $2R_{\text{half}}$. The last column presents the stellar, gas, and dark matter mass resolution. We separate the table into two parts. The upper part presents the different model galaxies at fixed fiducial yield set and the lower part presents the two galaxies rerun with different yield sets.

Simulation	Yield set	M_{star} ($10^{10} M_{\odot}$)	M_{gas} ($10^{10} M_{\odot}$)	R_{half} (kpc)	$\log(Z/Z_{\odot})$	$[\text{Fe}/\text{H}]$	$12 + \log(\text{O}/\text{H})$	$m_{\text{star}}/m_{\text{gas}}/m_{\text{dark}}$ ($10^5 M_{\odot}$)
galaxy models								
g2.79e12	new fid	18.23	21.70	2.81	−1.62	0.17	9.24	1.06/3.18/17.35
g8.26e11	new fid	4.00	8.00	2.93	−1.74	−0.06	9.01	1.06/3.18/17.35
g7.55e11	new fid	3.75	6.81	2.71	−1.77	−0.09	8.97	1.06/3.18/17.35
g2.19e11	new fid	0.08	0.92	2.78	−2.72	−1.05	8.02	0.13/0.39/2.17
g1.57e11	new fid	0.10	1.00	4.67	−2.78	−1.12	8.06	0.13/0.39/2.17
g4.99e10	new fid	0.01	0.19	2.93	−3.18	−1.51	7.53	0.04/0.11/0.64
g2.83e10	new fid	0.003	0.10	1.77	−3.39	−1.75	7.35	0.04/0.11/0.64
g7.05e09	new fid	0.0002	0.01	0.45	−3.51	−1.85	7.29	0.01/0.03/0.19
Model variations								
g8.26e11	alt	4.10	7.51	2.67	−1.80	−0.14	8.99	1.06/3.18/17.35
g8.26e11	alt2	3.52	7.69	2.61	−1.95	−0.17	8.74	1.06/3.18/17.35
g8.26e11	alt3	4.22	7.80	2.69	−1.74	−0.13	9.04	1.06/3.18/17.35
g8.26e11	alt4	3.91	7.65	2.68	−1.81	−0.17	9.04	1.06/3.18/17.35
g8.26e11	steepIMF	4.29	7.93	2.47	−2.01	−0.22	8.61	1.06/3.18/17.35
g8.26e11	longDelay	4.19	7.62	2.68	−1.73	−0.11	9.05	1.06/3.18/17.35
g8.26e11	highNorm	4.29	7.71	2.64	−1.71	0.03	9.06	1.06/3.18/17.35
g2.83e10	alt	0.004	0.12	1.76	−3.34	−1.79	7.47	0.04/0.11/0.64
g2.83e10	alt2	0.003	0.14	1.63	−3.57	−1.81	7.02	0.04/0.11/0.64
g2.83e10	alt3	0.003	0.13	2.18	−3.27	−1.74	7.52	0.04/0.11/0.64
g2.83e10	alt4	0.003	0.12	1.99	−3.49	−1.85	7.45	0.04/0.11/0.64
g2.83e10	longDelay	0.003	0.10	2.04	−3.35	−1.77	7.34	0.04/0.11/0.64
g2.83e10	highNorm	0.003	0.11	1.75	−3.33	−1.64	7.35	0.04/0.11/0.64
g2.83e10	steepIMF	0.006	0.09	1.67	−3.35	−1.59	7.24	0.04/0.11/0.64
g2.83e10	low fb	0.005	0.16	2.02	−3.19	−1.56	7.51	0.04/0.11/0.64

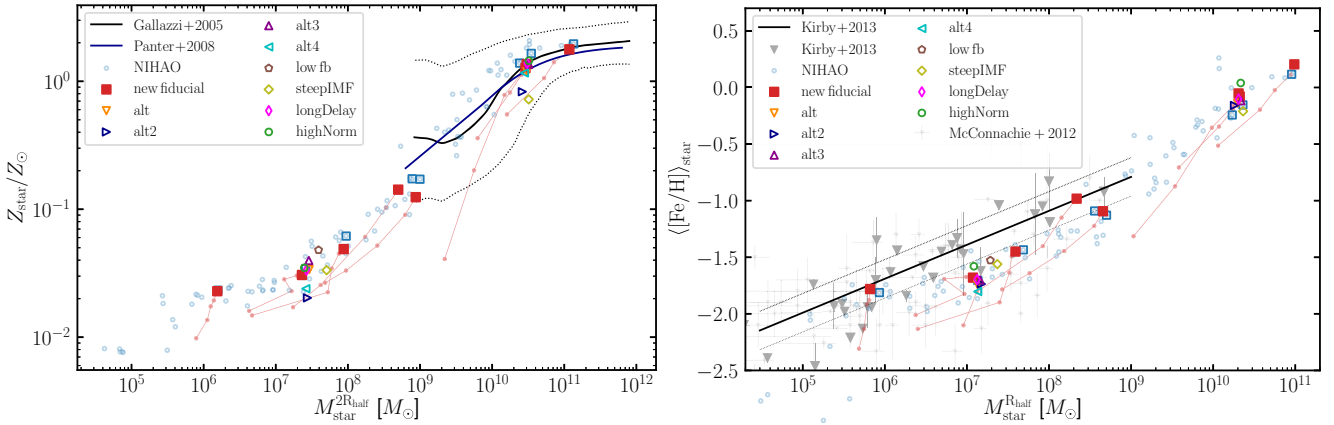


Figure 3. Stellar mass metallicity relation (left-hand panel) and stellar mass versus stellar iron abundance (right-hand panel) of the new chemical enrichment implemented in *Gasoline2* (red points) in comparison to observational data from Gallazzi et al. (2005), Panter et al. (2008), and Kirby et al. (2013), respectively. In the right-hand panel faint grey points show additional observational data from McConnachie (2012). We highlight the temporal evolution of our model galaxies with thin red dots connected with lines. Faint blue data points show results from the NIHAO sample highlighting the galaxies selected for re-simulation with blue squares. The open coloured symbols show additional runs with varying yield tables or stellar physics parameters as detailed in Table 2.

Let us first compare the results of the new chemical enrichment (red squares) with their corresponding NIHAO galaxies (blue open squares). We see that for almost all galaxies stellar masses, total metallicities and stellar iron abundances agree well between new and old runs. Deviations between individual galaxies are within the expected scatter introduced by stochastic star formation recipes (see also Genel et al. 2019; Keller et al. 2019). Note, the agreement between stellar masses measured at two different radii further implies

that the new chemical enrichment does not lead to drastic changes in stellar mass profiles. Comparing this further with the runs where we tested different yield sets from the literature (open coloured symbols), we see that the latter have a modest influence on the total metallicity or iron abundance but less so on the total stellar mass. We have checked that this differences in metallicities are solely due to different yields and not to a different star formation history (see Fig. B1 in the appendix).

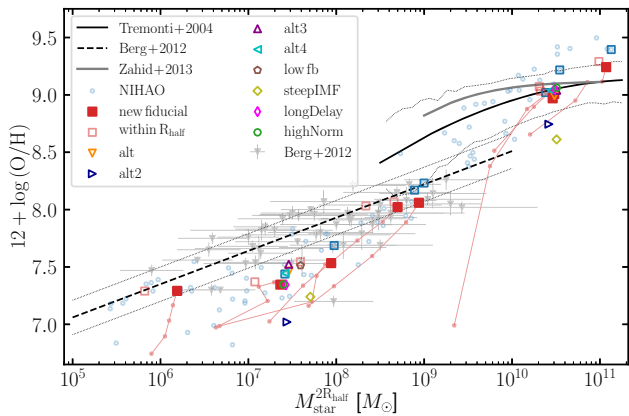


Figure 4. Stellar mass gas phase oxygen abundance of the new chemical enrichment implemented in *Gasoline2* (red points) in comparison with observational data from Tremonti et al. (2004), Zahid et al. (2013), and Berg et al. (2012), respectively. Again, we highlight the temporal evolution of our model galaxies with thin red dots connected with lines. The faint blue data points show results from the NIHAO simulations where we highlight the galaxies selected for re-simulation with blue squares. The open coloured symbols show additional runs with varying yield tables or stellar physics parameters as detailed in Table 2.

Let us now turn to compare the simulation results to the observed relations. For the total metallicity observed relations exist for the stellar mass range 10^9 – $10^{12} M_{\odot}$ while for the iron abundance Local Group dwarfs constrain the relation in the mass range $\sim 10^4$ – $10^9 M_{\odot}$. While the total metallicity of our simulations agrees well with the observed mass–metallicity relation, the iron abundance around $\sim 10^7 M_{\odot}$ falls slightly below the observed relation although within the scatter of individual observed dwarf galaxies. This is especially true for both the old NIHAO implementation and the new chemical enrichment model regardless of the yield table used. Note, we highlight the temporal evolution of our model galaxies with thin red dots connected with lines. This shows that the galaxies mostly evolve along the redshift zero relation.

4.2 Stellar mass versus gas oxygen abundance

Fig. 4 compares the gas phase oxygen abundance of our simulations to the observed relation of Tremonti et al. (2004), Zahid et al. (2013), and Berg et al. (2012). At the high mass end, the new chemical enrichment implementation results in lower oxygen abundances compared to the old NIHAO simulations. Our new results are in better agreement with the observations. At the low mass end we find no differences in the oxygen abundance between the different implementations and both agree with the observed relation. At intermediate stellar masses of 10^6 – $10^8 M_{\odot}$ we find similar results as for the stellar iron abundance. Both the new chemical enrichment model and the old NIHAO implementation underpredict the gas phase oxygen abundance. We have tested if this is affected by the spatial selection applied ($2R_{\text{half}}$ versus $1R_{\text{half}}$ shown in faint red open squares) but find only a mild dependence of the oxygen abundance on the spatial selection. However, the stellar mass measured for a smaller apertures shifts towards lower masses which slightly alleviates the tension between simulations and observations (especially for the models with larger O abundance like e.g. *alt*, *alt3*).

Finally, we find that only a change in the strength of stellar feedback (here a reduction of the coupling efficiency of the early stellar feedback from 13 per cent to 6 per cent, magenta open

pentagon) or a larger high mass slope of the IMF (olive open diamond) and as such a reduced amount of CC-SN tend to increase the iron abundance. This brings the simulations closer to the observed relations at intermediate stellar masses. Although we find that a steeper IMF high mass slope in combination with our fiducial yield table underpredicts the gas oxygen abundance stronger than the fiducial model.

Agertz et al. (2020) find that explosive feedback expels too many metals from low mass dwarf galaxies such that those feedback mechanisms result in a lower metallicity at fixed stellar mass while other dwarf galaxy scaling relations are unaffected by this. Our tests here seem to confirm their findings although a reduced feedback strength also results in roughly a factor of two larger stellar masses for the dwarf galaxy tested here. A larger suite of model galaxies at this mass scale is certainly needed to investigate this issue further and potentially constrain feedback mechanisms at this galaxy mass range. Also, care must be taken in comparing simulation results of isolated galaxies with relations obtained for Local Group galaxies which evidently live in a galactic environment influenced by the two major galaxies MW and Andromeda. This might bias Local Group galaxies which might have experienced accretion of pre-enriched gas or suffered some kind of stripping (e.g. Buck et al. 2019a) by ram pressure and tidal forces.

4.3 Metallicity and α -element distributions

Section 4.1 revealed that the average stellar metallicity of the simulations is in agreement with the observed mass–metallicity relation. However, it is further instructive to not only investigate zero-order moments of the abundance distribution but look at the distribution functions. In the next two subsections, we investigate in detail the metallicity distribution functions and the α element distribution functions for different yield sets probed in this work. Contrary to the previous section, we do not apply any spatial selection and use all star particles within the virial radius of the galaxies.

4.3.1 Metallicity distribution

Fig. 5 shows the distribution of stellar iron abundance for the dwarf galaxy $g2.83e10$ (left-hand column) and the MW mass galaxy $g8.26e11$ (right-hand column) for all eight different models. Fig. C1 in the appendix shows the distribution function of total metallicity for the two galaxies. In all panels the filled blue histogram shows the fiducial NIHAO implementation while in the left-hand panels in Fig. 5 the vertical thin dotted lines show the range of mean iron abundance of five dwarf galaxy satellites of M31 as measured by Kirby et al. (2019, table 4) and the black thin histograms in the right-hand panels show results from the DR3 of the GALAH survey based on a magnitude-limited selection and thus only selecting stars in the solar neighbourhood.

The distribution functions of all models share a common feature: They are skewed towards low metallicities/iron abundance. All models calculated for this work show a rather extended tail of low metallicity stars reminiscent of the well-known G dwarf problem (e.g. Pagel 1997). This common feature is a manifestation of the star formation history which follows the same shape in all eight models (see fig. B1). The maxima of our distribution functions of $[\text{Fe}/\text{H}]$ for the dwarf galaxy (left-hand panels of Fig. 5) agree well with the measured limits by Kirby et al. (2019) for all our models. However, the maxima of *longDelay*, *alt3*, and NIHAO are close to the observed upper limit. We furthermore see significant secondary peaks of *highNorm* and *steepIMF* above the limit.

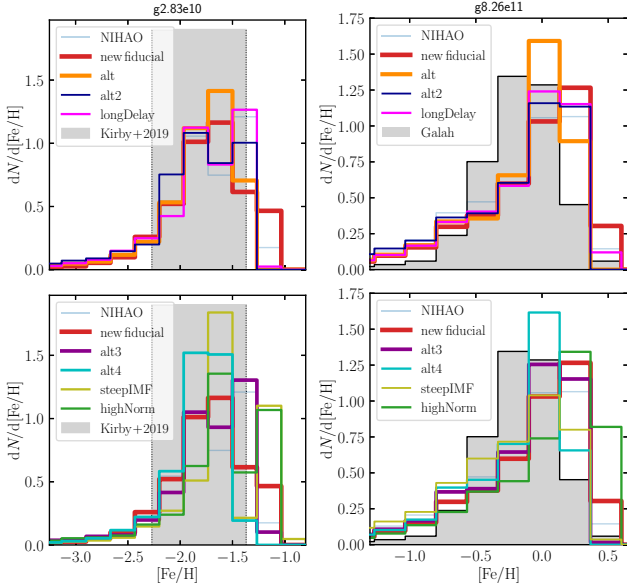


Figure 5. Distribution functions of stellar iron abundance for the dwarf galaxy $g2.83e10$ (left-hand panels) and the MW mass galaxy $g8.26e11$ (right-hand panels) for various different yield sets tested in this work. For comparison, in the left-hand panels, the grey bar indicates the range of mean iron abundance of the five dwarf satellites of M31 as measured by Kirby et al. (2019, table 4). In the right-hand panels, the grey filled histogram indicates the observed iron abundance distribution of solar neighbourhood stars from GALAH-DR3 (Buder et al. 2021). The light blue histograms show the results from the fiducial NIHAO implementation.

The shape of our distribution functions for the MW mass galaxy (right-hand panels of Fig. 5) agree qualitatively with the one measured by GALAH DR3 for the local MW, that is, a maximum around solar $[\text{Fe}/\text{H}]$ with a skewed shape, extending towards lower $[\text{Fe}/\text{H}]$. Quantitatively, we do see, however, that the measured distribution of GALAH DR3 peaks slightly below zero, i.e. $[\text{Fe}/\text{H}] \sim -0.1$ dex, similar to most studies of the Solar vicinity (e.g. Casagrande et al. 2011; Katz et al. 2021). Furthermore, the observed distribution is less skewed in GALAH DR3 compared to our implementations. Selecting stars only from a solar neighbourhood in the simulations does not change this conclusion as this effects both the lowest and highest metallicity bins. Whether this discrepancy is due to a different SFH in the simulations and the MW or due to different SN Ia yields needs to be investigated further.

In this work, we did not vary the yield set for SN Ia only their delay time and normalization and keep the yieldset fixed to Seitzzahl et al. (2013). For a more in depth exploration of the impact of different type Ia SN yields on MW chemical evolution see the recent work by Palla (2021). This work shows that Fe masses and $[\alpha/\text{Fe}]$ abundance patterns are less affected by the detailed yield sets while for Fe-peak elements (which are not the focus of this paper) there seems to be a strong dependence on the white dwarf explosion mechanism. Here we find that a difference in delay time between our fiducial model ($\tau = 40$ Myr) and the *longDelay* runs ($\tau = 160$ Myr) does not lead to significant differences in the iron abundance. However, a slight change in the SN Ia normalization (-2.9 versus -2.75 for the *highNorm* run) leads to a slight increase in the iron abundance for both the dwarf and the MW mass galaxy.

As previously noted, all models exhibit similar mean metallicities/iron abundances. Note, we do not recalibrate any simulation parameters in this work. This is important to note again as in the

next subsection we turn to investigate the distribution of α -elements relative to iron, the $[\alpha/\text{Fe}]$ ratio. Therefore we like to stress here that especially the iron abundance distributions of our models do not differ significantly and all differences discussed in the next section can be attributed to differences in yield tables (remember the similarity in star formation history).

4.3.2 Diversity in α -element distribution

The left-hand panels of Fig. 6 show the distribution of stellar α -element abundance for all eight simulations of $g2.83e10$ in comparison to observational data from Kirby et al. (2019, table 4) for five Andromeda satellite galaxies shown with the grey shaded area. The right-hand panels of Fig. 6 show the stellar α -element abundance for the eight models of the MW mass galaxy $g8.26e11$ in comparison to the GALAH DR3 data. Here we define the α -element abundance as $[\alpha/\text{Fe}] \equiv (\text{Mg} + \text{Si} + \text{Ca} + \text{Ti})/\text{Fe}$, in line with the definition of Kirby et al. (2019) who uses the union of Mg, Si, Ca, and Ti and GALAH DR3 which uses an uncertainty-weighted average of Mg, Si, Ca, and Ti. In contrast to the iron abundance distribution functions shown in Fig. 5 the α -element distribution shows significant diversity between different yield tables or feedback strengths. For the dwarf galaxy the peaks of the different α -element distributions span a range of 0.5 dex between $[\alpha/\text{Fe}] \sim -0.25$ (*alt2*) up to $[\alpha/\text{Fe}] \gtrsim 0.3$ (*alt4*). Comparing to the observations by Kirby et al. (2019) of dwarf galaxies of similar stellar mass puts strong constraints on valid yield sets. Only the alternative yield set *alt* and *alt3* result in $[\alpha/\text{Fe}]$ ratios in agreement with observed α -enhancements. All other models underpredict the observed values except for the yield set *alt4* which strongly overproduces α -elements. We have checked that differences in stellar mass between our model galaxies and the observed dwarfs do not influence our conclusions. The left-hand panels of Fig. C2 in the appendix show that the average α -element abundance is not sensitive to the stellar mass of the dwarfs and only shows a mild dependence on stellar mass. This implies that cosmic variance that manifests itself with different star formation histories for the dwarf galaxies does not have a strong impact on the α -element abundance distribution at constant stellar mass. This opens the possibility to constrain chemical enrichment models by comparing those results to accurate and correct measurements of stellar masses and abundances.

Finally, we find that most of the differences of the different models are caused by differences in the Mg abundances as shown in Fig. C3 in the appendix.

Note that the differences in α -element abundance is not caused by a difference in iron abundance as we have discussed in the previous section. Especially the *steepIMF* and *alt2* model which underpredict the α -element abundance show iron abundances in the same range as the other models. Similarly in order to explain the overproduction of α -elements in the *alt4* model by a miss-match in the iron abundance we would expect this model to underproduce iron compared to the other models. This is clearly not the case as the lower right-hand panel of Fig. 5 shows.

For the MW mass galaxy (right-hand panels of Fig. 6) the differences are not as pronounced as for the dwarf but variations among the models are still quite noticeable. The peaks of the different α -element distributions span a range of ~ 0.3 dex between $[\alpha/\text{Fe}] \sim -0.15$ (*steepIMF*) up to $[\alpha/\text{Fe}] \gtrsim 0.2$ (*alt3*). Qualitative comparison to observational data from the GALAH survey (grey histograms) suggest $[\alpha/\text{Fe}]$ values around solar hinting at

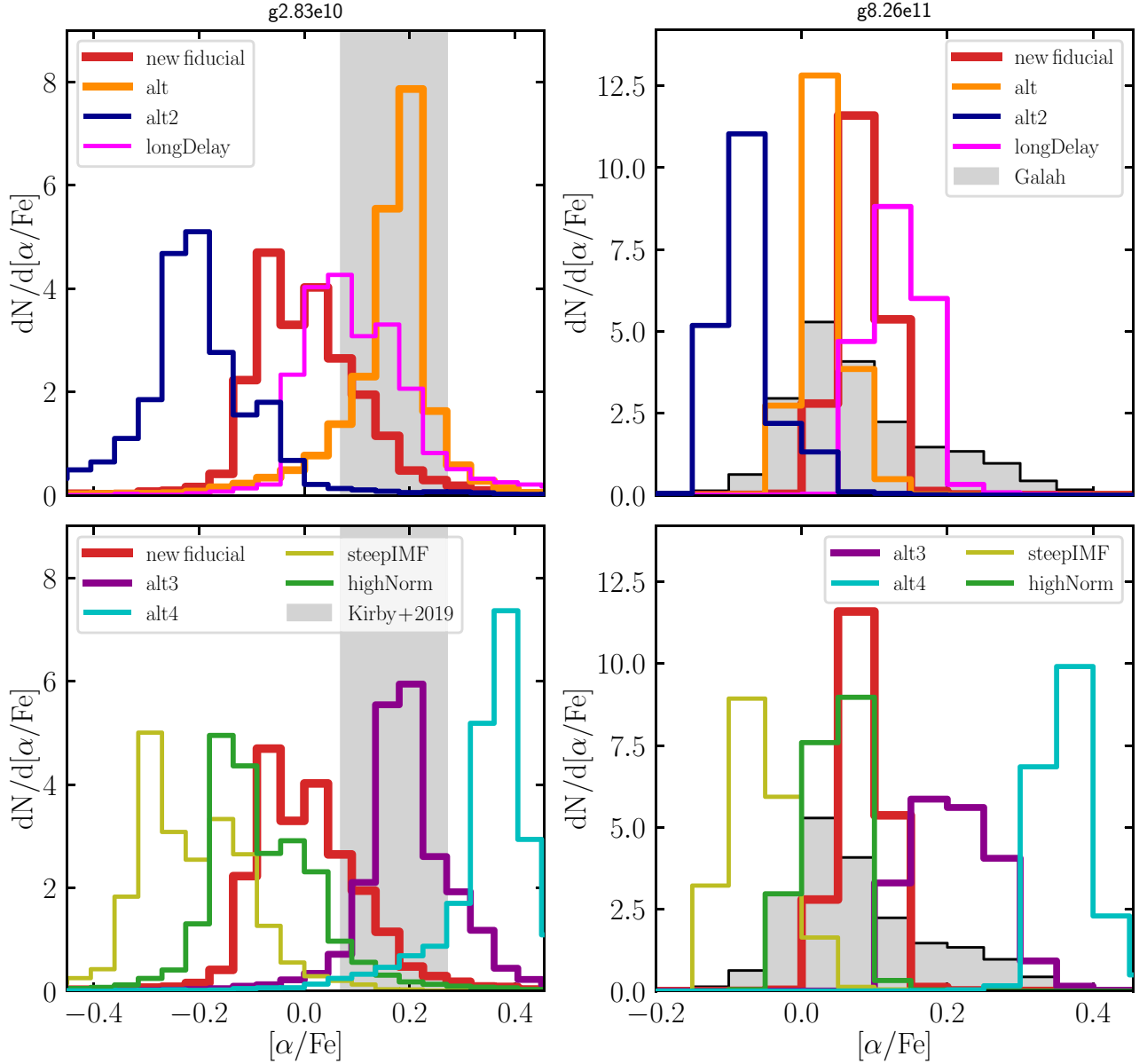


Figure 6. Distribution functions of α -element abundance ($[\alpha/\text{Fe}] \equiv (\text{Mg} + \text{Si} + \text{Ca} + \text{Ti})/\text{Fe}$) for the dwarf galaxy $g2.83e10$ (left-hand panels) and the MW mass galaxy $g8.26e11$ (right-hand panels) for various different yield sets tested in this work. In the left-hand panels, the grey vertical bar shows the mean α -abundance of five dwarf galaxy satellites of M31 as measured by Kirby et al. (2019, table 4). In the right-hand panels, the grey/black histograms show results for the solar neighbourhood from the GALAH DR3 (Buder et al. 2021).

the conclusion that the *longDelay* and *alt*, *alt3*, *alt4* yield sets overpredict the α -enhancement while the *steepIMF* and *alt2* model tend to underproduce the α -enhancement. Only the models *fiducial* and *highNorm* are marginally in agreement with the GALAH data (see also Fig. 7). This finding is at odds with the results for the dwarf galaxy where the best matching models are the *alt*, *alt3*, and *longDelay* yield sets. Note, the exact value of the α -enhancement does not depend strongly on the stellar mass of the galaxy as the right-hand side of Fig. C2 in the appendix shows.

Note, the comparison to GALAH is only of qualitative nature for several reasons. First, the GALAH survey selects for primarily nearby main sequence stars near the Sun. Our comparison is not entirely consistent given as we select stars across a much larger radial extent in the simulation and not sampled as per the GALAH survey.

However, we have checked that only selecting solar neighbourhood stars does not affect our results in the right-hand panels of Fig. 6 and subsequent conclusions, as this only affects the edges of the alpha-distributions at very low and high α -abundances. Secondly, we are comparing a generic MW mass model to MW data and it is a priori unclear if the particular model galaxy is a good MW analogue. Thirdly, MW represents a single object and it is still unclear if the observed α -enhancement is representative for all MW mass galaxies. Current theoretical models do not agree on this point and offer explanations for both scenarios (see e.g. Grand et al. 2018; Mackereth et al. 2018; Clarke et al. 2019; Buck 2020; Renaud et al. 2021b). However, a quantitative comparison needs to take into account the exact selection function of the particular survey and account for measurement uncertainty on stellar abundances and metallicities.

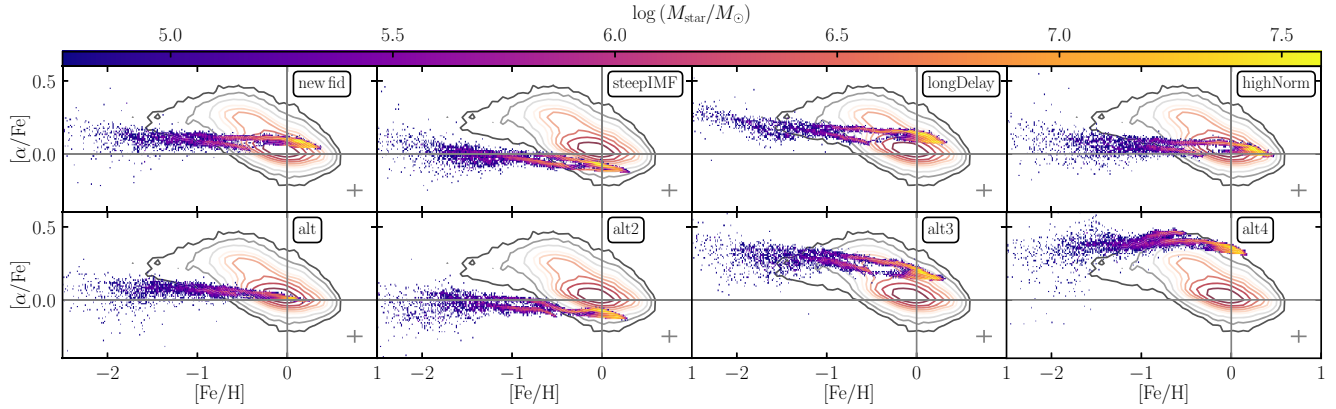


Figure 7. Tinsley–Wallerstein diagram: Distribution of α -element abundances as a function of metallicity, coloured by stellar mass. For the simulation we show only stellar particles within a radial range of $6.5 < R < 9.5$ kpc and absolute vertical height of $|z| < 2$ kpc from the disc mid-plane. Colour-coding shows the stellar mass in each $[\alpha/\text{Fe}], [\text{Fe}/\text{H}]$ bin. For comparison, contours show stellar number counts for the solar neighbourhood from GALAH DR3 (Buder et al. 2021). The grey error bars in the lower right of each panel indicate the median observational uncertainty of the GALAH data. The grey lines at $[\text{Fe}/\text{H}] = 0$ and $[\alpha/\text{Fe}]$ indicate solar abundance values. See the text for more explanation.

Therefore we defer a more thorough comparison to a dedicated future study.

Comparing the distribution functions of α -elements for the dwarf galaxy in the left-hand side of Fig. 6 with those for the MW mass galaxy, we find for all yield models a consistent shift of ~ 0.1 dex to larger values of $[\alpha/\text{Fe}]$ for the more massive galaxy. This is also consistent for the other galaxies run with the fiducial yield table in our sample (see Fig. C2 in the appendix). On the other hand, for the observations, to zero order we find that the median $[\alpha/\text{Fe}]$ decreases with stellar mass when going from dwarf to MW mass scale as might be expected from a larger contribution of SN Ia to the enrichment in MW mass galaxies and varying star formation efficiencies in the dwarf galaxies (e.g. Nidever et al. 2020). Of course, this is a simplified comparison as also in the dwarf regime we expect the α -abundance to be a function of metallicity (e.g. Fig. 8 Kirby et al. 2019). However, understanding the exact cause of this finding and how much of it can be attributed to different star formation efficiencies in the simulations and the observations needs a larger sample of dwarf galaxies and a more in depth analysis which is outside the scope of this paper which focuses on different yield tables. Future work is needed in order to gauge if this finding poses a challenge to our current models and if so, researching the reason for this opposite trend of α -enhancement with stellar mass in the simulations and the observations will open new possibilities for our understanding of star formation and feedback in (dwarf) galaxies.

4.4 Abundance tracks in $[\text{X}/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$

It is now well known that for the MW the evolution of the α -abundance as a function of metallicity shows a very distinct trend (e.g. Wallerstein 1962; Tinsley 1979; Wyse & Gilmore 1995; Fuhrmann 1998; Venn et al. 2004; Bensby et al. 2014; Hayden et al. 2015). Low metallicity *in-situ* stars follow a plateau at high α -abundances which obeys a knee at slightly subsolar metallicities and bends towards lower α -abundance at higher metallicity. Simultaneously there exist a lower α -abundance sequence around solar α -abundance across a wide range of metallicities. This peculiar tracks have been explained by various authors to arise from an exquisite interplay of star formation time-scales, stellar yields, and particular gas accretion history of the MW (e.g. Chiappini et al. 1997; Matteucci 2012; Spitoni et al. 2020). Therefore this plane of observables offers an

ideal constraint for the interplay of yield sets and feedback model. In Buck (2020), we have shown that such a bimodal sequence is a rather generic feature of the NIHAO feedback scheme. In this section, we use this finding to scrutinize different yield sets in their ability to reproduce (at fixed feedback parameters) the location and shape of the observed α -abundance bimodality. Here two things are of importance: On the one hand, the enrichment to high α -abundances of ~ 0.3 dex at low metallicity and simultaneously solar values at solar metallicity should be achieved. And secondly, the shape/position of the knee/down bending should occur roughly at the observed metallicity as it constrains the turning point where the relative contribution of CC-SN to SN Ia changes.

In Fig. 7, we compare the $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ plane for all our models. The upper left-hand panel shows the fiducial yield set and to the right of it we compare the three models varying physical parameters of the SSP while the bottom row shows the four alternative yield sets tested. In order to guide the eye, we indicate solar values with grey lines and overlay observational results from the GALAH survey. Note, for this work we apply only a simple spatial selection to the simulated data and refrain from modelling in detail the GALAH selection function (e.g. accounting for that fact that GALAH mainly observes cool dwarfs). Additionally, we accompany this plot by Figs 8 and 9 which compare single element (columns) abundance tracks among all models (rows). For this first exploratory work, we have decided to focus on a differential comparison between yield sets. Therefore we refrain from exactly modelling the GALAH selection function and decided to simply select from the simulation star particles within an annulus of $6.5 < R < 9.5$ kpc. This spatial selection approximates the solar-centric distance distribution of stars included in GALAH. In the simulation we are formally dealing with star particles several orders of magnitude more massive than single stars. Therefore, for the simulation we show a 2d histogram of stellar mass while for the GALAH survey the contours show number counts. In future works one can improve upon this simple selection by generating synthetic stellar catalogues from the simulations (e.g. Valentini et al. 2018), and applying the exact GALAH selection function to these catalogues.

In all panels, two clear sequences for the simulations can be seen: an extended primary sequence with the highest stellar mass densities reaching supersolar $[\text{Fe}/\text{H}]$ and showing a bimodality at the highest metallicities, and a secondary sequence at low $[\text{Fe}/\text{H}]$ as

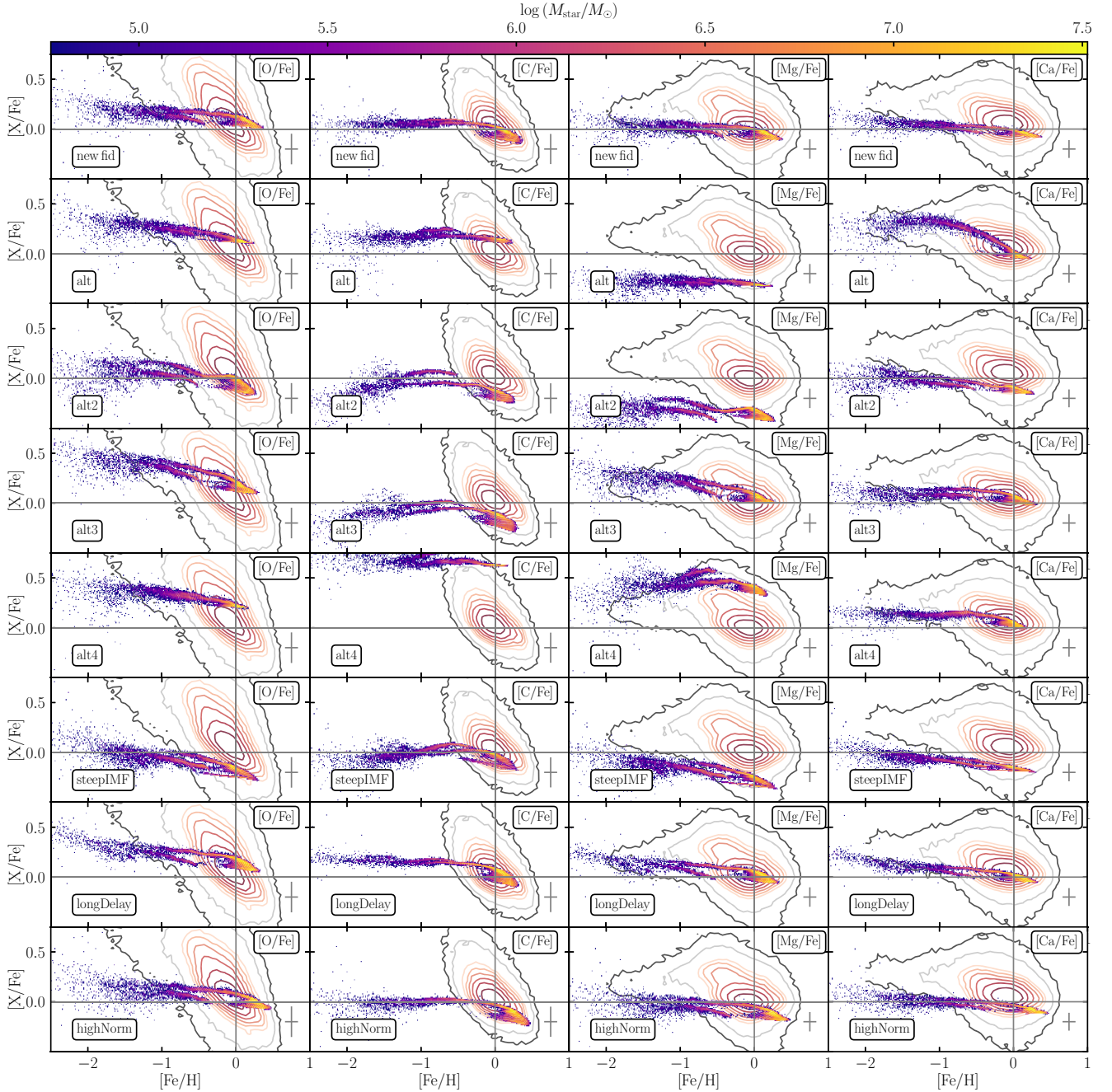


Figure 8. Distribution of O, C, Mg, and Ca (from left to right) abundances as a function of metallicity, coloured by stellar mass. We show only stellar particles within $6.5 < R < 9.5$ kpc and $|z| < 2$ kpc. Different rows compare different yield sets as indicated in the lower left-hand corner of each panel. For comparison, contours show the results of the GALAH survey (Buder et al. 2021). The grey error bars in the lower right of each panel indicate the median observational uncertainty of the GALAH data and grey lines indicate solar abundance values. Note, for the *alt4* yield set the C abundance extends to outside the plotting range with $[C/Fe] \sim 0.8$ (see also Fig. C4).

well as low $[\alpha/Fe]$ (Fig 7). Note, our choice of contour lines for the GALAH observations makes it harder to see the bimodality in the observed data. The primary sequence are *in-situ* stars of the main galaxy, whereas the secondary sequence originates from accreted satellite galaxies in these simulations (see Buck et al. 2020, for more details). Recent literature (Hayes et al. 2018; Das et al. 2020; Buder et al. 2021) have found similarly pronounced differences between accreted and *in-situ* stars also for the MW. Their lower abundances in α -elements as well as Na and Al agree

qualitatively with our study. However, our predictions for $[C/Fe]$ are above the primary sequence, whereas they have found to be below the primary sequence in observations (e.g. Hayes et al. 2020) found $[C/Fe]$ to be below the primary sequence at the same $[Fe/H]$ for the MW.

Subsequently, we first focus on the primary sequence, before addressing the secondary sequence.

Fig. 7 confirms our previous notion of distinctive differences among different yield sets. We find a range of shapes for the α -

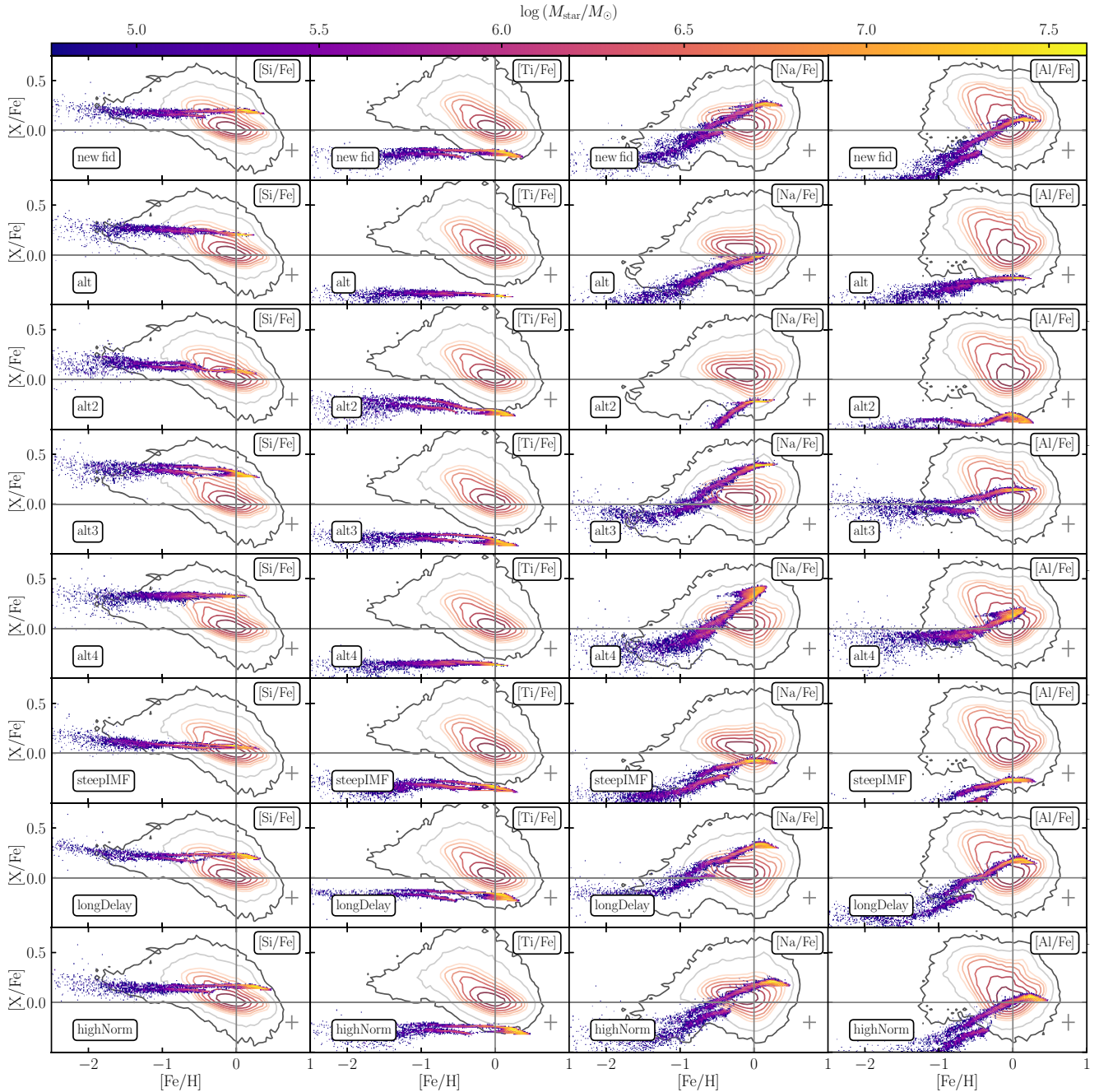


Figure 9. Same as Fig. 8 but this time for Si, Ti, Na, and Al (from left to right).

bimodality ranging from a single sequence evolving from high to low $[\alpha/\text{Fe}]$ (*alt*) over a less separated bimodality (*fiducial model*) up to well separated sequences (all other models). Especially the amount of α -enhancement differs between different yield sets as would be expected from their different predictions for the element release (e.g. Fig. 2).

Notably the *steepIMF* model which uses the same yield set as the fiducial model heavily underproduces the α -abundance at all metallicities. This is simply caused by a reduction of the number of high mass stars which contribute most to the α -element synthesis. Contrary, the *alt3* model overproduces α -elements and never reaches solar values despite a clear bimodality. The strongest separation of the two sequences can be seen in the *longDelay* model which shifts

the onset of SNIa to ~ 160 Myr which is a factor of 4 longer delay time compared to the fiducial model. The *highNorm* run fits the overall $[\alpha/\text{Fe}]$ values of the MW better than the fiducial run but the larger number of SN Ia leads to slightly too large iron abundance for the low- α stars.

Finally, the double knee seen in model *alt2* is due to a slightly modified SFH in this model which leads to lower values of the peak SFR between redshift $z = 2-1$ ($\sim 3-5$ Gyr after the big bang) as can be seen from Fig. B1 in the appendix.

Let us turn to compare the intrinsic scatter seen in the simulation data to the observed broad sequences. In each panel the grey cross in the lower right corner highlights the median observational uncertainty of the GALAH abundances. While accurately forward

modelling the simulation data to account for the exact survey selection function and the observational uncertainties is outside the scope of this work we still recognize that the simulated sequences exhibit a narrower intrinsic scatter. Comparing those narrow tracks with the median observational uncertainty (the grey cross) it seems that even when we would account for the broad observational uncertainty in future work the simulated tracks are still too narrow. This poses a challenge for our theoretical models to explain chemical enrichment in MW mass galaxies.

Figs 8 and 9 add a more diverse view to this as different elements X obey different slopes in the $[X/Fe]$ versus $[Fe/H]$ plane both in the models and in the GALAH data. Especially the very steep slope observed for O is not reproduced by any of our models.⁸ C is relatively similar for all models with less variation compared to O or Mg for example. Similar to C, Ca tracks are relatively similar in all our models except for the model *alt* which shows way stronger Ca enhancement at low metallicities than all other models. For Si, Na and Al we find similar strong differences among yield sets. Only for Ti those differences are less pronounced. In fact, Si is almost always overproduced while Ti is underproduced in all models which is a known problem of the theoretical models (see e.g. discussion of fig. 23 in Kobayashi et al. 2020) which might be resolved in 2 or 3D models of nucleosynthesis. Finally, the trend of $[Na/Fe]$ from our simulations agrees qualitatively well with low $[Na/Fe]$ for metal poor stars. The disc values for $[Na/Fe]$, however, show a plateau-like distribution within the data, but a steep increase for all our model realizations (with exception of the *steepIMF*, *longDelay* and *highNorm* realizations showing a plateau, but at a wrong $[Na/Fe]$ level). The comparison between data and simulation for Al is especially difficult, because the observational data are limited by the detectability of Al lines in the metal-poor regime. The trend of $[Al/Fe]$ is however expected to be very similar to $[Na/Fe]$ also for metal-poor stars. The exact reason for the tension of $[Na/Fe]$ and $[Al/Fe]$ for metal-rich stars is still elusive and further research is necessary.

Comparing the scatter of the model galaxies with the observations shows that also for individual elements the simulations reveal quite narrow tracks. For example, the Mg track of the fiducial model agrees quite well with the centroid of the GALAH data but when taking into account the median observational uncertainty of ~ 0.05 dex the simulated data would still be narrower than the observations. Reasons for this miss-match might be that the observations cover a large range of signal-to-noise ratios and thus uncertainties (Buder et al. 2021) as well as that observational uncertainties might be underestimated for parts of the parameter space. For example, the boutique analysis of the Mg abundance from only highest quality spectra of the AMBRE project finds a much reduced intrinsic scatter compared to GALAH (Santos-Peral et al. 2020), but covers only a few thousand stars and has a less well described selection function than GALAH. Future observations and re-analysis of stellar spectra might decrease the observational scatter and allow to better constrain the intrinsic abundance scatter in the MW (see also Bedell et al. 2018; Sharma et al. 2020; Yuxi et al. 2021, for a determination of the intrinsic abundance scatter in the MW). Another reason might certainly be an underestimation of the intrinsic scatter in the simulations.

⁸It is still debated, if this steep trend is the true astrophysical one, or introduced by the observational analysis. E.g. APOGEE measures a shallower trend in the IR compared to optical studies (see e.g. fig. 22 in Buder et al. 2018). The cause for this discrepancy is still not clear and explanations range from atomic data or 3D NLTE effects in the IR to other not explained effect in the optical.

Several model simplifications might artificially decrease the intrinsic scatter: A simplistic feedback prescription might fail to reproduce the intrinsic scatter observed for the MW. Assuming a fully sampled IMF might in fact drive the chemical enrichment to the mean of an SSP neglecting the effects of stochasticity on the enrichment process. Similarly, overestimating the effect of turbulent metal mixing (see Section 3.1 for details on our implementation) in the ISM or underestimating the effect of radial migration decreases the intrinsic abundance scatter at fixed radius. The effects of turbulent metal mixing are studied in detail by Escala et al. (2018) who find that sub-grid turbulent metal mixing strongly reduces the scatter in abundance distributions. Similar conclusions are also made by Rennehan (2021) for anisotropic mixing. Thus, the observed abundance scatter allows us to constrain valid sub-grid metal mixing procedures. We will do so in future work.

The primary focus of this work is to study different yield sets and evaluate their viability in reproducing the observational constraints. Therefore we did not model observational uncertainties in the simulation as this would have hindered our comparison. However, future work will address this point and model observational selection effects and uncertainties. Exactly pin-pointing if there is indeed a discrepancy between simulation results and MW observations as well as researching the exact cause for it, may it be observational or theoretical, is certainly needed and will help us understand galaxy formation and the MW in more depth.

5 CONCLUSION AND OUTLOOK

5.1 Summary and conclusion

By coupling a flexible chemical evolution/stellar population model with a well tested galaxy formation model, we are not only able to study in more depth physical processes of galaxy formation, but more importantly, we are able to constrain stellar evolution and elemental synthesis models. Our chemical evolution model allows us to flexibly explore stellar evolution parameters in a full cosmological galaxy formation model. This helps us to put constraints on valid parameter combinations by comparing simulation results to chemical abundance structures as observed in the MW and local galaxies. In particular, we show that at fixed SSP parameters (i.e. stellar lifetimes, mass ranges for CC-SN and SN Ia, etc.) different yield sets from the literature show significant difference in their ability to enrich the galaxies with different chemical species. Especially the α -element abundances of different yield sets show significant differences in dwarf and MW mass galaxies. This in combination with exquisite observational data for the MW and its dwarf galaxies could potentially be used to discriminate between different models. Similarly there exist difference in the chemical abundances for slight variations of the SN Ia delay time or the high mass slope of the IMF.

Our results are summarized as follows:

(i) We have presented a new, mass- and metallicity-dependent, time-resolved chemical enrichment model for cosmological simulations that is able to follow any element evolution self-consistently. This model enables a flexible parameter combination with the potential of parameter inference from observational data via the Chempy package prior to the simulation run, i.e. optimization with a computational cheap model.

(ii) We implement our new model in the *Gasoline2* code within the NIHAO feedback model. Figs 3 and 4 show that this new model reproduces stellar mass metallicity and gas phase oxygen abundance relations generally well. We find a slight tension of the models around

10^7 – $10^8 M_{\odot}$. Investigating the reason for this requires a larger sample of galaxies at these stellar masses but it might point towards a failure of the current feedback model at those stellar masses as previously shown by Agertz et al. (2020). Future work will investigate this point in more depth.

(iii) We find that total metallicity and iron abundances (Figs 3 and 4) as well as their distribution functions (Figs 5 and C1) do not differ much between different yield sets when used with the NIHAO feedback model. We find that for all yield sets tested no recalibration of the feedback model is necessary in order to match observational constraints on the mass–metallicity relations.

(iv) The α -element distribution function (Fig. 6) as well as the $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ plane (Fig. 7) show significant diversity between the models and can help putting constraints on valid yield set combinations or SSP parameter choices such as SN Ia delay time or IMF shape/slopes, when compared with observational data, as done in Figs 7, 8, and 9.

(v) We find that none of our models is able to simultaneously match observational constraints for dwarf galaxies and the MW when looking at the α element distribution function (Fig. 6). While in the dwarf galaxy regime the models *alt*, *alt3* and *longDelay* seem to best match the (Kirby et al. 2019) data for the MW the best matching models are the *fiducial* and *highNorm* models. While most of our simulated α element distribution functions show an increasing median with stellar mass, the observations are consistent with the opposite trend, thus signalling an unique *diversity of abundance patterns*, which we find difficult to reconcile in a single chemical enrichment model.

(vi) For the MW model we find a narrow intrinsic scatter around the tracks in the $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ plane (Fig. 7) and similarly for individual elements (Figs 8 and 9) which compared to the observational scatter is in slight tension with MW observations, the *problem of matching the scatter in chemical abundance space*. Possible solutions to this might be an underestimation of stochasticity from the assumption of a fully sampled IMF and a possible underestimation of observational uncertainties. Another uncertainty is given by the sub-grid metal mixing prescription in the simulation. However, for a quantitative conclusion a more in depth analysis of the simulations including the exact modelling of the observational selection function is needed and will be done in future work.

(vii) Within the framework of the NIHAO feedback model the best parameter combination we have found are tables by Karakas & Lugaro (2016) for AGB stars with tables from Chieffi & Limongi (2004) for CC-SN and yields from Seitenzahl et al. (2013) for SN Ia with a power law delay time distribution function following Maoz, Mannucci & Brandt (2012) with a delay time of ~ 40 Myr and normalization of -2.9 . Those are combined with stellar lifetimes from Argast et al. (2000) and a fiducial Chabrier (2003) IMF where the maximum mass of stars exploding as CC-SN is set to $40 M_{\odot}$ above which we assume direct black hole collapse. The minimum stellar mass exploding as CC-SN is set to $8 M_{\odot}$, see Table A1 for more details. A nearly identical parameter combination has been inferred from MW data previously by Philcox et al. (2018). Note, for the dwarf galaxy we have found that either a steeper IMF or a higher SN Ia normalization brings its iron abundance closer to the observed relation of Kirby et al. (2013). This might be an indication that a variable IMF model might actually be preferable.

The combination of a generic parameter inference tool with fully cosmological galaxy simulations can help us improve our understanding of galaxy formation in general and the chemo-kinematics

of our own MW in particular. We are confident that our approach of flexibly varying essential stellar evolution and chemical enrichment parameters will positively contribute to our understanding of galaxies and might further help us to improve nucleosynthetic yields and feedback prescriptions.

5.2 Outlook

The chemical enrichment prescription presented in this paper offers several new pathways of studying galaxy formation across cosmic time. Fig. 1 shows the relative contribution of different sources (AGB stars, SN Ia, and CC-SN) to the abundance of different chemical elements. Specifically studying those elements which result predominantly from a single source might help us gain deeper insight into the (astro)physical mechanisms of galaxy formation and evolution.

Our new model further offers the possibility to study the relative contribution of different enrichment channels in L^* galaxies (SNII versus AGB versus SNIa) which is not possible from observational data alone. Similarly the MW is the only galaxy to date for which we have quantitative resolved spectroscopic observations of single stars to study its chemokinematics. Assessing to which extent the features we see in the MW are a generic outcome of galaxy formation or the result of a particular formation history is challenging to answer. Studying the chemokinematic properties of a larger ensemble of simulated L^* galaxies where the survey specific selection functions are taken into account will certainly boost our knowledge. Therefore a more in depth and quantitative comparison to MW data will be performed in a follow-up study.

At the dwarf galaxy scale, we find opposite trends with stellar mass for the median $[\alpha/\text{Fe}]$ values in simulations and observations. Certainly the star formation efficiency in dwarf galaxies plays a large role in setting their α -enhancement. Whether the α -element distributions of simulated dwarf galaxies, in fact, in disagreement with the observed diversity of abundance tracks in Local Group dwarf galaxies remains to be seen. Therefore future work will thoroughly compare of model dwarf galaxy abundance tracks with observed element distributions and help us understand the physical processes of galaxy formation in more depth.

ACKNOWLEDGEMENTS

TB thanks Martin Rey, Cristina Chiappini, Oscar Agertz (and his research group) as well as Andrey Kravtsov (and his research group) for insightful discussions on an earlier version of this draft. TB and CP acknowledge support from the European Research Council under ERC-CoG grant CRAGSMAN-646955. JR was funded by the DLR (German space agency) via grant 50 QG 1403. AO is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 443044596. This research made use of PYNBODY (Pontzen et al. 2013), SCIPY (Virtanen et al. 2020), NUMPY (Harris et al. 2020), IPYTHON (Pérez & Granger 2007) and JUPYTER (Kluyver et al. 2016) and used MATPLOTLIB (Hunter 2007) to display all figures in this work. We gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (www.lrz.de). This research was carried out on the High Performance Computing resources at New York University Abu Dhabi. We greatly appreciate the contributions of all these computing allocations. All simulations performed for this study including test runs and failed simulations used about 2.5M core-hours which under the assumption of 30W per core and a CO_2

intensity of electricity of $\sim 600 \text{ kgCO}_2 \text{ MWh}^{-1}$ amounts to a carbon footprint of $\sim 45 \text{ t}$ of CO_2 .

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. Python scripts and data to reproduce all plots of this paper can be found at: https://github.com/TobiBu/chemical_enrichment.git.

REFERENCES

- Agertz O., Teyssier R., Moore B., 2009, *MNRAS*, 397, L64
- Agertz O. et al., 2020, *MNRAS*, 491, 1656
- Agertz O. et al., 2021, *MNRAS*, 503, 5826
- Anders F. et al., 2014, *A&A*, 564, A115
- Argast D., Samland M., Gerhard O. E., Thielemann F. K., 2000, *A&A*, 356, 873
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Baade W., 1944, *ApJ*, 100, 137
- Bedell M. et al., 2018, *ApJ*, 865, 68
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Bensby T., Feltzing S., Oey M. S., 2014, *A&A*, 562, A71
- Berg D. A. et al., 2012, *ApJ*, 754, 98
- Bertschinger E., 2001, *ApJS*, 137, 1
- Blancato K., Ness M., Johnston K. V., Rybizki J., Bedell M., 2019, *ApJ*, 883, 34
- Boeche C. et al., 2011, *AJ*, 142, 193
- Boeche C. et al., 2013, *A&A*, 553, A19
- Brook C. B., Stinson G., Gibson B. K., Shen S., Macciò A. V., Obreja A., Wadsley J., Quinn T., 2014, *MNRAS*, 443, 3809
- Brook C. B., Kawata D., Gibson B. K., Gallart C., Vicente A., 2020, *MNRAS*, 495, 2645
- Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, *ApJ*, 655, L17
- Buck T., 2020, *MNRAS*, 491, 5435
- Buck T., Macciò A. V., Obreja A., Dutton A. A., Domínguez-Tenreiro R., Granato G. L., 2017, *MNRAS*, 468, 3628
- Buck T., Ness M. K., Macciò A. V., Obreja A., Dutton A. A., 2018, *ApJ*, 861, 88
- Buck T., Macciò A. V., Dutton A. A., Obreja A., Frings J., 2019a, *MNRAS*, 483, 1314
- Buck T., Dutton A. A., Macciò A. V., 2019b, *MNRAS*, 486, 1481
- Buck T., Ness M., Obreja A., Macciò A. V., Dutton A. A., 2019c, *ApJ*, 874, 67
- Buck T., Obreja A., Macciò A. V., Minchev I., Dutton A. A., Ostriker J. P., 2020, *MNRAS*, 491, 3461
- Buder S. et al., 2018, *MNRAS*, 478, 4513
- Buder S. et al., 2019, *A&A*, 624, A19
- Buder S. et al., 2021, *MNRAS*, 506, 150
- Casagrande L., Schönrich R., Asplund M., Cassisi S., Ramírez I., Meléndez J., Bensby T., Feltzing S., 2011, *A&A*, 530, A138
- Chabrier G., 2001, *ApJ*, 554, 1274
- Chabrier G., 2003, *PASP*, 115, 763
- Chiappini C., Matteucci F., Gratton R., 1997, *ApJ*, 477, 765
- Chiappini C., Romano D., Matteucci F., 2003, *MNRAS*, 339, 63
- Chieffi A., Limongi M., 2004, *ApJ*, 608, 405
- Clarke A. J. et al., 2019, *MNRAS*, 484, 3476
- Côté B., West C., Heger A., Ritter C., O'Shea B. W., Herwig F., Travaglio C., Bisterzo S., 2016a, *MNRAS*, 463, 3755
- Côté B., Ritter C., O'Shea B. W., Herwig F., Pignatari M., Jones S., Fryer C. L., 2016b, *ApJ*, 824, 82
- Das P., Hawkins K., Jofré P., 2020, *MNRAS*, 493, 5195
- Dayal P., Ferrara A., Dunlop J. S., 2013, *MNRAS*, 430, 2891
- Dehnen W., Aly H., 2012, *MNRAS*, 425, 1068
- Doherty C. L., Gil-Pons P., Lau H. H. B., Lattanzio J. C., Siess L., 2014, *MNRAS*, 437, 195
- Doherty C. L., Gil-Pons P., Siess L., Lattanzio J. C., 2017, *Publ. Astron. Soc. Aust.*, 34, e056
- Doménech-Moral M., Martínez-Serrano F. J., Domínguez-Tenreiro R., Serna A., 2012, *MNRAS*, 421, 2510
- Domínguez-Tenreiro R., Obreja A., Brook C. B., Martínez-Serrano F. J., Serna A., 2017, *ApJ*, 846, 72
- Dutton A. A. et al., 2017, *MNRAS*, 467, 4937
- Dutton A. A., Macciò A. V., Buck T., Dixon K. L., Blank M., Obreja A., 2019, *MNRAS*, 486, 655
- Dutton A. A., Buck T., Macciò A. V., Dixon K. L., Blank M., Obreja A., 2020, *MNRAS*, 499, 2648
- Eilers A.-C., Hogg D. W., Rix H.-W., Frankel N., Hunt J. A. S., Fouvry J.-B., Buck T., 2020, *ApJ*, 900, 186
- Escala I. et al., 2018, *MNRAS*, 474, 2194
- Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761
- Feuillet D. K., Bovy J., Holtzman J., Girardi L., MacDonald N., Majewski S. R., Nidever D. L., 2016, *ApJ*, 817, 40
- Few C. G., Courty S., Gibson B. K., Kawata D., Calura F., Teyssier R., 2012, *MNRAS*, 424, L11
- Fishlock C. K., Karakas A. I., Lugaro M., Yong D., 2014, *ApJ*, 797, 44
- François P., Matteucci F., Cayrel R., Spite M., Spite F., Chiappini C., 2004, *A&A*, 421, 613
- Freeman K., Bland-Hawthorn J., 2002, *ARA&A*, 40, 487
- Frischknecht U. et al., 2016, *MNRAS*, 456, 1803
- Fryer C. L., 1999, *ApJ*, 522, 413
- Fuhrmann K., 1998, *A&A*, 338, 161
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, 362, 41
- Genel S. et al., 2019, *ApJ*, 871, 21
- Gibson B. K., 1997, *MNRAS*, 290, 471
- Gibson B. K., Loewenstein M., Mushotzky R. F., 1997, *MNRAS*, 290, 623
- Gibson B. K., Pilkington K., Brook C. B., Stinson G. S., Bailin J., 2013, *A&A*, 554, A47
- Grand R. J. J. et al., 2018, *MNRAS*, 474, 3629
- Gutcke T. A., Springel V., 2019, *MNRAS*, 482, 118
- Haardt F., Madau P., 2012, *ApJ*, 746, 125
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Hayden M. R. et al., 2015, *ApJ*, 808, 132
- Hayes C. R. et al., 2018, *ApJ*, 852, 49
- Hayes C. R. et al., 2020, *ApJ*, 889, 63
- Hilmi T. et al., 2020, *MNRAS*, 497, 933
- Hirschmann M., De Lucia G., Fontanot F., 2016, *MNRAS*, 461, 1760
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Iwamoto K., Brachwitz F., Nomoto K., Kishimoto N., Umeda H., Hix W. R., Thielemann F.-K., 1999, *ApJS*, 125, 439
- Karakas A. I., 2010, *MNRAS*, 403, 1413
- Karakas A. I., Lugaro M., 2016, *ApJ*, 825, 26
- Katz D., Gomez A., Haywood M., Snaith O., Di Matteo P., 2021, preprint ([arXiv:e-prints](https://arxiv.org/abs/2108.00013))
- Kawata D., 2001, *ApJ*, 558, 598
- Kawata D., Gibson B. K., 2003, *MNRAS*, 340, 908
- Keller B. W., Wadsley J. W., Wang L., Kruijssen J. M. D., 2019, *MNRAS*, 482, 2244
- Kennicutt R. C., Jr., 1998, *ARA&A*, 36, 189
- Khoperskov S., Haywood M., Snaith O., Di Matteo P., Lehnert M., Vasiliev E., Naroenkov S., Berczik P., 2021, *MNRAS*, 501, 5176
- Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, *ApJ*, 779, 102
- Kirby E., Gilbert K. M., Wojno J., Escala I., 2019, American Astronomical Society Meeting Abstracts #233, 316.05
- Kluyver T. et al., 2016, in Loizides F., Schmidt B., eds, Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS Press, the Netherlands, p. 87. Available at: <https://eprints.soton.ac.uk/403913/>
- Kobayashi C., Karakas A. I., Lugaro M., 2020, *ApJ*, 900, 179
- Kravtsov A. V., 2013, *ApJ*, 764, L31

- Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, *Astron. Lett.*, 44, 8
- Kroupa P., Tout C. A., Gilmore G., 1993, *MNRAS*, 262, 545
- Lia C., Portinari L., Carraro G., 2002, *MNRAS*, 330, 821
- Limongi M., Chieffi A., 2018, *ApJS*, 237, 13
- Mackereth J. T. et al., 2017, *MNRAS*, 471, 3057
- Mackereth J. T., Crain R. A., Schiavon R. P., Schaye J., Theuns T., Schaller M., 2018, *MNRAS*, 477, 5072
- Maiolino R., Mannucci F., 2019, *A&AR*, 27, 3
- Maoz D., Mannucci F., 2012, *Publ. Astron. Soc. Aust.*, 29, 447
- Maoz D., Sharon K., Gal-Yam A., 2010, *ApJ*, 722, 1879
- Maoz D., Mannucci F., Brandt T. D., 2012, *MNRAS*, 426, 3282
- Martig P., 2001, *A&A*, 370, 194
- Martig M. et al., 2016, *MNRAS*, 456, 3655
- Martínez-Serrano F. J., Serna A., Domínguez-Tenreiro R., Mollá M., 2008, *MNRAS*, 388, 39
- Matteucci F., 2012, *Chemical Evolution of Galaxies*. Springer, Berlin
- McConnachie A. W., 2012, *AJ*, 144, 4
- Miller G. E., Scalo J. M., 1979, *ApJS*, 41, 513
- Minchev I., Chiappini C., Martig M., 2013, *A&A*, 558, A9
- Miranda M. S. et al., 2016, *A&A*, 587, A10
- Mo H., van den Bosch F. C., White S., 2010, *Galaxy Formation and Evolution*. Cambridge Univ. Press, Cambridge
- Mollá M., Díaz A. I., 2005, *MNRAS*, 358, 521
- Mosconi M. B., Tissera P. B., Lambas D. G., Cora S. A., 2001, *MNRAS*, 325, 34
- Moster B. P., Naab T., White S. D. M., 2018, *MNRAS*, 477, 1822
- Müller B., 2016, *Publ. Astron. Soc. Aust.*, 33, e048
- Naiman J. P. et al., 2018, *MNRAS*, 477, 1206
- Ness M. K., Johnston K. V., Blancato K., Rix H. W., Beane A., Bird J. C., Hawkins K., 2019, *ApJ*, 883, 177
- Nidever D. L. et al., 2020, *ApJ*, 895, 88
- Nomoto K., Kobayashi C., Tominaga N., 2013, *ARA&A*, 51, 457
- Obreja A., Stinson G. S., Dutton A. A., Macciò A. V., Wang L., Kang X., 2016, *MNRAS*, 459, 467
- Obreja A. et al., 2019a, *MNRAS*, 487, 4424
- Obreja A., Macciò A. V., Moster B., Udrescu S. M., Buck T., Kannan R., Dutton A. A., Blank M., 2019b, *MNRAS*, 490, 1518
- Oppenheimer B. D., Davé R., 2008, *MNRAS*, 387, 577
- Pagel B. E. J., 1997, *Nucleosynthesis and Chemical Evolution of Galaxies*. Cambridge Univ. Press, Cambridge
- Palla M., 2021, *MNRAS*, 503, 3216
- Panther B., Jimenez R., Heavens A. F., Charlot S., 2008, *MNRAS*, 391, 1117
- Penzo C., Macciò A. V., Casarini L., Stinson G. S., Wadsley J., 2014, *MNRAS*, 442, 176
- Pérez F., Granger B. E., 2007, *Comput. Sci. Eng.*, 9, 21
- Philcox O., Rybizki J., Gutcke T. A., 2018, *ApJ*, 861, 40
- Pignatari M. et al., 2016, *ApJS*, 225, 24
- Pillepich A. et al., 2018, *MNRAS*, 473, 4077
- Planck Collaboration VIII, 2014, *A&A*, 571, A16
- Pontzen A., Roškar R., Stinson G. S., Woods R., Reed D. M., Coles J., Quinn T. R., 2013, *pynbody: Astrophysics Simulation Analysis for Python*
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, 334, 505
- Prantzos N., Abia C., Limongi M., Chieffi A., Cristallo S., 2018, *MNRAS*, 476, 3432
- Price D. J., 2008, *J. Comput. Phys.*, 227, 10040
- Raiteri C. M., Villata M., Navarro J. F., 1996, *A&A*, 315, 105
- Recchi S., Matteucci F., D'Ercole A., 2001, *MNRAS*, 322, 800
- Renaud F., Agertz O., Read J. I., Ryde N., Andersson E. P., Bensby T., Rey M. P., Feuillet D. K., 2021a, *MNRAS*, 503, 5846
- Renaud F., Agertz O., Andersson E. P., Read J. I., Ryde N., Bensby T., Rey M. P., Feuillet D. K., 2021b, *MNRAS*, 503, 5868
- Rennehan D., 2021, *MNRAS*, 506, 2836
- Richings A. J., Schaye J., 2016, *MNRAS*, 458, 270
- 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
- Romano D., Karakas A. I., Tosi M., Matteucci F., 2010, *A&A*, 522, A32
- Rybizki J., Just A., Rix H.-W., 2017, *A&A*, 605, A59
- Saitoh T. R., 2017, *AJ*, 153, 85
- Saitoh T. R., Makino J., 2009, *ApJ*, 697, L99
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Santos-Peral P., Recio-Blanco A., de Laverny P., Fernández-Alvar E., Orde-novic C., 2020, *A&A*, 639, A140
- Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2005, *MNRAS*, 364, 552
- Seitenzahl I. R. et al., 2013, *MNRAS*, 429, 1156
- Sestito F. et al., 2021, *MNRAS*, 500, 3750
- Sharma S. et al., 2020, preprint ([arXiv:2011.13818](https://arxiv.org/abs/2011.13818))
- Shen S., Wadsley J., Stinson G., 2010, *MNRAS*, 407, 1581
- Sim S. A. et al., 2013, *MNRAS*, 436, 333
- Spitoni E., Vincenzo F., Matteucci F., 2017, *A&A*, 599, A6
- Spitoni E., Verma K., Silva Aguirre V., Calura F., 2020, *A&A*, 635, A58
- Starkenburg E., Shetrone M. D., McConnachie A. W., Venn K. A., 2014, *MNRAS*, 441, 1217
- Steinmetz M., Mueller E., 1994, *A&A*, 281, L97
- Steinmetz M. et al., 2020, *AJ*, 160, 83
- Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, *MNRAS*, 373, 1074
- Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, *MNRAS*, 428, 129
- Thielemann F. K. et al., 2003, *Nucl. Phys. A*, 718, 139
- Tinsley B. M., 1979, *ApJ*, 229, 1046
- Tissera P. B., White S. D. M., Scannapieco C., 2012, *MNRAS*, 420, 255
- Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, *MNRAS*, 382, 1050
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Truelove J. K., Klein R. I., McKee C. F., Holliman II J. H., Howell L. H., Greenough J. A., 1997, *ApJ*, 489, L179
- Valentini M., Bressan A., Borgani S., Murante G., Girardi L., Tornatore L., 2018, *MNRAS*, 480, 722
- Valentini M., Borgani S., Bressan A., Murante G., Tornatore L., Monaco P., 2019, *MNRAS*, 485, 1384
- van Dokkum P. G. et al., 2013, *ApJ*, 771, L35
- Venn K. A., Irwin M., Shetrone M. D., Tout C. A., Hill V., Tolstoy E., 2004, *AJ*, 128, 1177
- Ventura P., Di Criscienzo M., Carini R., D'Antona F., 2013, *MNRAS*, 431, 3642
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, *Nat. Rev. Phys.*, 2, 42
- Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, *MNRAS*, 387, 427
- Wadsley J. W., Keller B. W., Quinn T. R., 2017, *MNRAS*, 471, 2357
- Wallerstein G., 1962, *ApJS*, 6, 407
- Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, *MNRAS*, 454, 83
- Weinberg D. H., Andrews B. H., Freudenburg J., 2017, *ApJ*, 837, 183
- Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, *MNRAS*, 399, 574
- Wyse R. F. G., Gilmore G., 1995, *AJ*, 110, 2771
- Xiang M. et al., 2019, *ApJS*, 245, 34
- Yuxi, Lu Ness M., Buck T., Zinn J., 2021, preprint ([arXiv:2102.12003](https://arxiv.org/abs/2102.12003))
- Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, *ApJ*, 771, L19

APPENDIX A: EXTENDED SIMPLE STELLAR POPULATION MODEL

A1 Stellar initial mass function

The amount, N , of stars of given stellar mass m included in the SSP are set by the functional form of the IMF which in the fiducial case

Table A1. Free stellar evolution parameters, θ , used in Chempy together with the fiducial values adopted in this work.

θ	Description	$\bar{\theta}_{\text{fiducial}}$
IMF type	functional form of IMF	
α_{IMF}	Chabrier high-mass slope	-2.3
	IMF mass range	0.1–100 M_{\odot}
	CC-SN mass range	8–40 M_{\odot}
	SN Ia delay time exponent	-1.12
$\log_{10}(N_{\text{Ia}})$	normalization of SN Ia rate	-2.9
τ_{Ia}	SN Ia delay time in Gyr	40 Myr
Z_{SSP}	metallicity of the SSP	$10^{-5}Z_{\odot} - 2Z_{\odot}$

is assumed to follow the Chabrier (2003) IMF

$$\frac{dN}{dm} = 0.158 \exp \left[-\frac{(\log m/M_{\odot} - \log 0.079)^2}{2 \times 0.69^2} \right] m < 1M_{\odot}$$

$$\frac{dN}{dm} = A (m/M_{\odot})^{-(1+\alpha_{\text{IMF}})} m > 1M_{\odot}, \quad (\text{A1})$$

where A is determined by a smooth transition of the two regimes and the high-mass slope (α_{IMF}) is one of Chempy's basic free parameters. For this work, we assume α_{IMF} to be universal with respect to stellar metallicity and cosmic time. The parameter α_{IMF} sets the ratio of low- to high-mass stars and controls the number of high-mass stars. Thus for our purposes the exact choice of α_{IMF} strongly influences the number of CC-SN and therefore the available feedback energy and elemental composition of the feedback. The total range of stellar masses over which the IMF extends can also be set and we have chosen a range from $m_{\text{min}} = 0.1$ to $m_{\text{max}} = 100M_{\odot}$ (see also Table A1).

Chempy offers the possibility to flexibly choose the functional form of the IMF. Next to the Chabrier (2003) IMF there are the Chabrier (2001), the Kroupa, Tout & Gilmore (1993), Salpeter (1955) and a generic broken power-law IMF (e.g. Miller & Scalo 1979) implemented. In the top panel of Fig. A1 we show a comparison of our fiducial IMF with a high mass slope of $\alpha_{\text{IMF}} = -2.3$ (blue line), its 2σ deviation $\alpha_{\text{IMF}} = -2.7$ (orange line Côté et al. 2016b), with the other IMFs implemented in Chempy.

A2 Stellar lifetimes

Newly synthesized elements are released into the ISM at the end of a star's lifetime either via AGB star winds or via SN explosions. In order to determine when those events take place we need to specify a stellar lifetime function which in general are a strong function of initial stellar mass and only weakly depend on stellar metallicity (see e.g. lower panel of Fig. A1). Chempy implements the stellar lifetime functions of Argast et al. (2000) and Raiteri et al. (1996) where the former is our fiducial choice. The lower panel of Fig. A1 shows stellar lifetimes as a function of stellar mass at solar metallicity for both parametrizations (blue solid and orange dashed line) in comparison to intrinsic lifetime variations predicted for different initial metallicities (red and green lines). This clearly shows that the metallicity dependence of stellar lifetimes is stronger than the differences between different authors.

Most massive stars will end their lives as a CC-SNe and feedback mass and energy into the ISM. The exact stellar mass which divides exploding from non-exploding stars is still under debate (e.g. Doherty et al. 2017) but usually assumed to be around $m_{\text{CC-SN}} = 8M_{\odot}$ (vertical grey dashed line in Fig. A1). Similarly there is evidence that the highest mass stars might end their lives with a direct collapse into

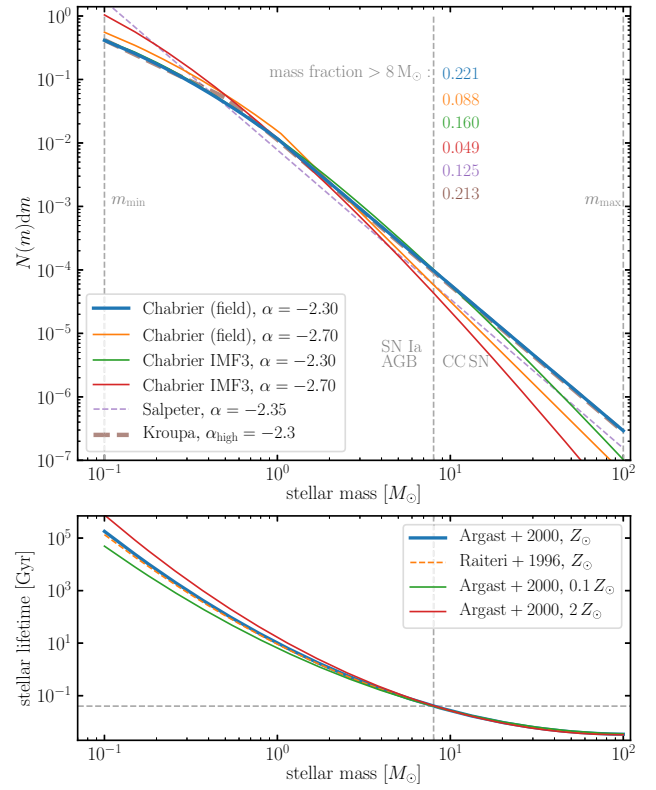


Figure A1. The simple stellar population model parameters: The stellar initial mass function (IMF, top panel) and the stellar lifetime function (lower panel). In the top panel, we compare different IMFs as indicated in the legend. For each IMF the coloured numbers indicate the mass fraction of star above $8M_{\odot}$ up to the assumed maximum mass of $100M_{\odot}$. The middle panel compares stellar lifetimes as a function of stellar mass for models from Argast et al. (2000) for three different metallicities with a model from Raiteri et al. (1996) at solar metallicity. The grey dashed lines indicate the lifetime of an $8M_{\odot}$ star which is ~ 40 Myr.

a black hole not releasing any energy or elements into the surrounding ISM. In this work, we assume that all stars more massive than $40M_{\odot}$ (e.g. Fryer 1999) will follow this path and set the maximum mass of stars that explode as CC-SN to $40M_{\odot}$.

Combining the IMF with the stellar lifetime function we are able to calculate the time evolution of CC-SN explosions and the number of stars in the AGB phase for a given SSP. In the upper panel of Fig. A2 we compare our fiducial model (Chabrier IMF) evolution to an alternative model using the Kroupa IMF. The most important difference is the number of high mass stars given by the different high mass slopes of the IMFs which manifests itself in a different number of CC-SN. The Chabrier IMF results in roughly twice the number of CC-SN compared to the Kroupa IMF.

A3 SN Ia delay times

The progenitor stars as well as the explosion mechanism resulting in SN Ia explosions is less well understood. Therefore we treat the fraction of stars and the mass range in which stars explode as SN Ia empirically. We parametrize the delay time distribution (DTD) of the SN Ia explosions with a power law following Maoz et al. (2010) (see also Table A1):

$$N_{\text{SN Ia}} \propto t^{-1.12}. \quad (\text{A2})$$

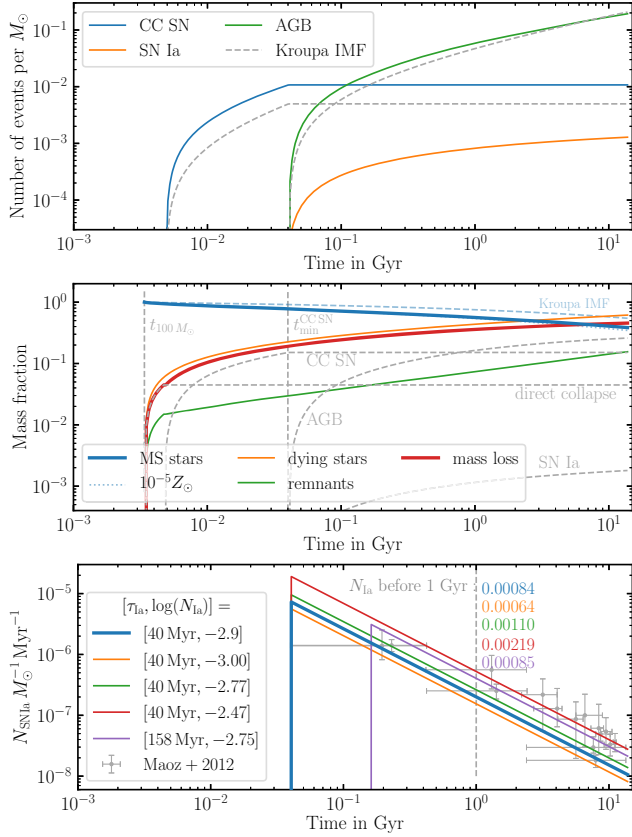


Figure A2. The upper panel shows the cumulative number of CC-SN, SN Ia, and AGB stars as a function of time at solar metallicity assuming a fiducial Chabrier IMF (solid coloured lines) while grey dashed lines assume a Kroupa IMF. The middle panel shows the corresponding cumulative mass fraction as a function of time of different components of the SSP. The solid blue line shows the mass fraction in main sequence stars as a function of SSP age assuming a Chabrier IMF while the dashed blue line assumes a Kroupa IMF. The red solid line shows the total mass loss of the population (assuming a Chabrier IMF) which is lower than the mass fraction of dying stars (orange line) by the amount of mass in stellar remnants (green line, e.g. stellar mass black holes or white dwarfs). The grey dashed lines show the mass-loss for each individual channel (CC-SN, AGB stars, and SN Ia). The two vertical dashed lines show the lifetime of a $100 M_{\odot}$ star and the lifetime of the minimum mass star exploding as CC-SN ($8 M_{\odot}$ here). The bottom panel shows the empirical SNIa delay time distribution (Maoz, Sharon & Gal-Yam 2010) with parameters taken from Maoz et al. (2012) and Maoz & Mannucci (2012). The grey data points show the observational data taken from Maoz et al. (2012).

The normalization, the number of SN Ia per solar mass over a time of 15 Gyr (N_{Ia}), and the time delay until the first SN Ia events occur (τ_{Ia}) are free parameters obtained from observations (e.g. Maoz et al. 2012). The bottom panel of Fig. A2 shows various realizations of the delay time distribution compared to the data on SN Ia explosions by Maoz et al. (2012). Our default model assumes the SN Ia normalization (N_{Ia}) from Maoz & Mannucci (2012, table 1) and the delay time (τ_{Ia}) is taken to be the minimum lifetime of stars that might explode as SN Ia, $\tau_{\text{Ia}} = 40$ Myr for $8 M_{\odot}$ stars. The resulting number of SN Ia explosions as a function of time is shown in Fig. A2. Through our empirical choice of parametrization this is independent of the exact IMF chosen.

We like to note here, that Chempy is able to calculate the distribution of stars along the IMF and the SN Ia explosions as a

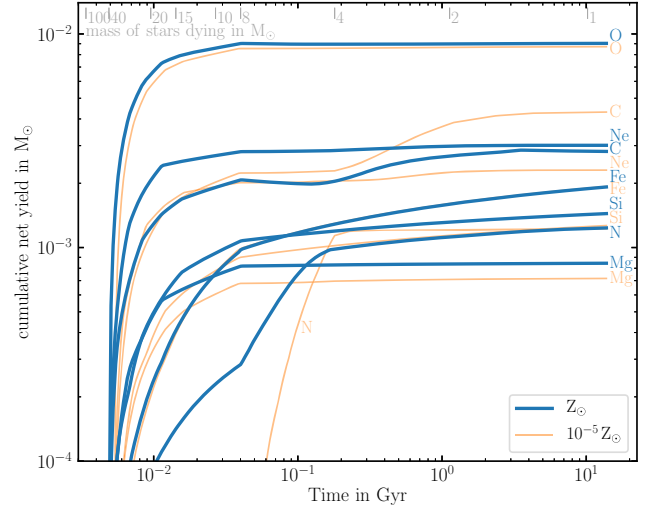


Figure A3. This figure shows the corresponding cumulative ejected SSP mass split into different elements for two different metallicities assuming stellar yields taken from Chieffi & Limongi (2004), Karakas & Lugaro (2016), Seitzzahl et al. (2013).

stochastic process. However, as soon as the SSP mass is sufficiently large ($\gtrsim 10^4 M_{\odot}$), this converges rapidly to the analytic solution. Therefore and for the sake of computational simplicity, we employed the faster analytic version in this work.

A4 Stellar mass-loss

We consider three different channels contributing to the mass and elemental enrichment: Supernova of type Ia (SN Ia), core-collapse supernova (CC-SNe) and asymptotic giant branch stars (AGB). For the latter two channels we consider metallicity and mass dependent enrichment, i.e. the elemental enrichment and the remnant mass depend on stellar metallicity as well as the mass of dying stars⁹ and as such explicitly on stellar lifetime. We further assume that all relevant materials are only ejected at the end of the star's lifetime. SN Ia on the other hand follow a different explosion path and we assume no metallicity dependence. The equations of how to calculate the mass-loss of a given SSP between times t_0 and t_1 are e.g. given in Section 4, equations (2) and (3) of Wiersma et al. (2009). The resulting time release of mass into the ISM and the accompanied mass-loss of the SSP particle is shown in the middle panel of Fig. A2. The blue solid line shows the evolving mass of main sequence stars in the SSP for our fiducial model (Chabrier IMF), the dashed blue line shows an alternative model assuming a Kroupa IMF. Comparing the two lines we see that the choice of the IMF has a significant impact on the expected mass-loss, less so the initial metallicity of the SSP which is shown with a blue dotted line and almost follows the fiducial solid blue line.

The thin orange line shows the amount of mass of dying stars as a function of time. Not all the mass of dying stars is returned to the ISM as the exploding stars leave behind a stellar remnant (green line, black hole or white dwarf depending on the star's mass). The red solid line thus shows the amount of mass returned to the ISM which results from subtracting the remnant mass from the mass of dying stars.

⁹The energy release of CC-SNe, on the other hand, is fixed to the fiducial value of 10^{51} erg per explosion and we do not consider energetic feedback from AGB winds for now.

Thin grey dashed lines split the mass loss into the different channels traced by our model. Note, after a Hubble time most mass returned to the ISM comes from AGB star winds followed by CC-SN. For the most massive stars undergoing direct black hole collapse, we assume a mass fraction of 0.75 is returned to the ISM. SN Ia explosions feed back only very little absolute mass. However, as we will see later their contribution to the iron mass is still significant.

APPENDIX B: STAR FORMATION HISTORY

Fig. B1 shows the cumulative mass growth history of the dwarf galaxy $g2.83e10$ (upper panel) and the MW mass galaxy $g8.26e11$ (lower panel) for all yield sets studied in this work. For the dwarf galaxy the *low fb* and the *steepIMF* runs lead to stronger mass growth in the early universe because of less feedback. This effect is not so much seen in the MW mass model.

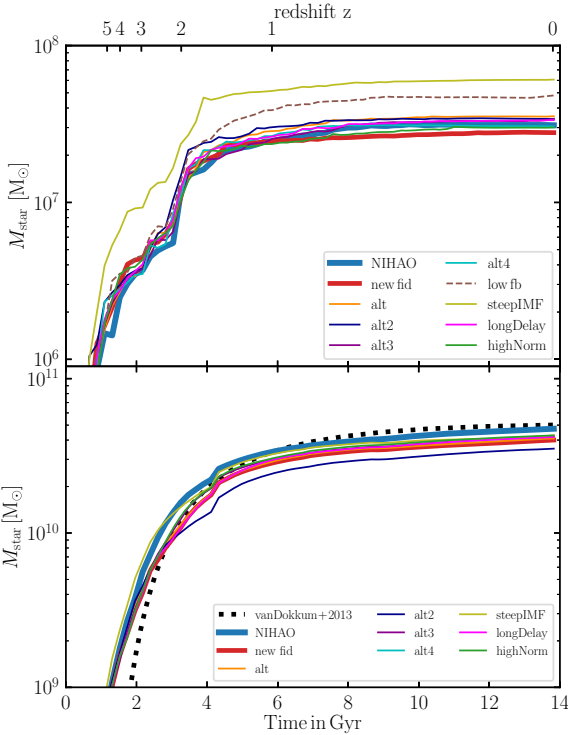


Figure B1. Stellar mass growth for $g2.83e10$ (upper panel) and $g8.26e11$ (lower panel) for all model runs performed for this work. The thick blue line shows the NIHAO simulation. In the lower panel the black dashed line shows the observationally inferred stellar mass growth for MW progenitor galaxies from van Dokkum et al. (2013). In order to enhance the slight difference at lower redshift, we cropped the plot at a lower mass limit of $10^9 M_\odot$.

APPENDIX C: METALLICITY DISTRIBUTION FUNCTION

Fig. C1 shows the distribution functions of total stellar metallicity for the dwarf galaxy $g2.83e10$ (left-hand panels) and the MW mass galaxy $g8.26e11$ (right-hand panels) for various different yield sets tested in this work.

Fig. C2 shows the α -element distribution function for our fiducial model comparing all model galaxies. The left-hand panel shows the three dwarf galaxies while the right-hand panel shows the more massive galaxies including the three L^* galaxies ($g2.79e12$, $g8.26e11$, $g7.55e11$). The left-hand panel shows the expected dependence of average α -enhancement on stellar mass for the dwarf galaxies where more massive galaxies ($g4.99e10$) show larger α -enhancement. The overall width of the distribution is rather independent of stellar mass. On the other hand, the more massive galaxies

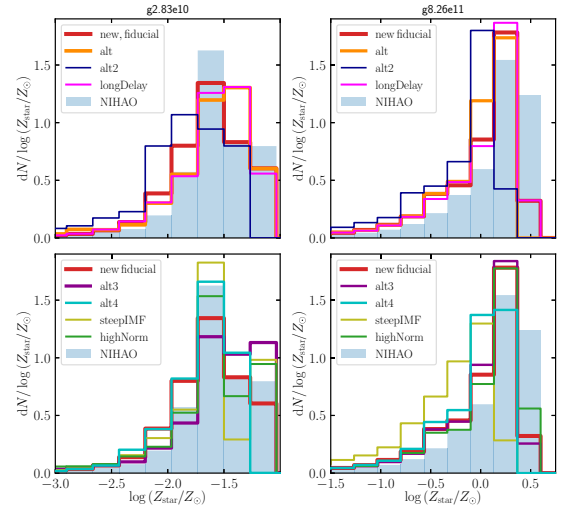


Figure C1. Distribution functions of total stellar metallicity for the dwarf galaxy $g2.83e10$ (left-hand panels) and the MW mass galaxy $g8.26e11$ (right-hand panels) for various different yield sets. The blue filled histograms show the results for the NIHAO model.

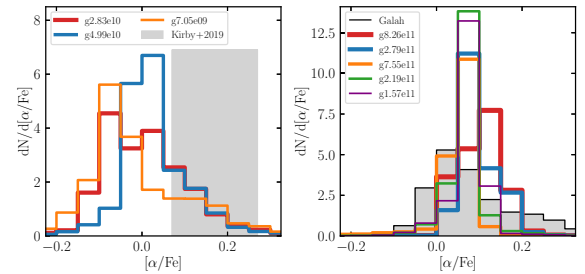


Figure C2. Distribution functions of α -element abundance for the dwarf galaxies (left-hand panel) and more massive galaxies (right-hand panel) for the fiducial yield set. The grey bar in the left-hand panel shows observational data from Kirby et al. (2019) and the filled histogram in the right-hand panel show the GALAH data.

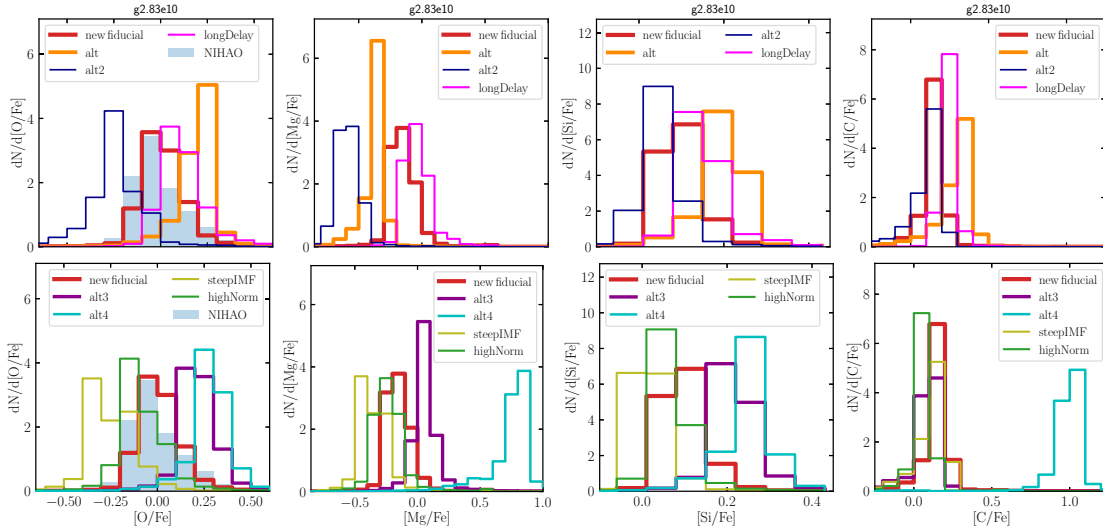


Figure C3. Distribution functions of O (left-hand panel), Mg (second column), Si (third column), and C (right-hand panel) for the dwarf galaxy g2.83e10 for various different yield sets tested in this work. The grey filled histogram in the left-hand panels show the results for the fiducial NIHAO implementation.

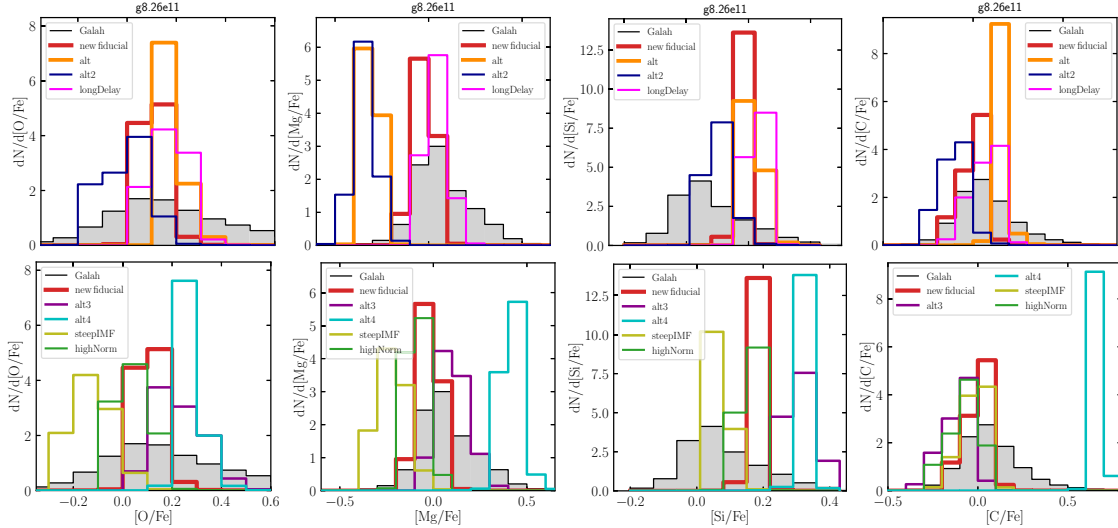


Figure C4. Same as Fig. C3 but for the MW mass galaxy g8.26e11 for various different yield sets tested in this work. The grey filled histogram shows observational data for the MW from the GALAH survey.

do not show such a dependence of average α -enhancement on stellar mass but they do show indications that the width of the distribution seems to be broader for the more massive galaxies (g2.79e12, g8.26e11). Figs C3 and C4 show additional distribution functions for the elements O, Mg, Si, and C.

APPENDIX D: ADDITIONAL CC SN YIELDS

Fig. D1 shows the same as Fig. 2 in the main text but for three additional CC-SN yield sets from the literature.

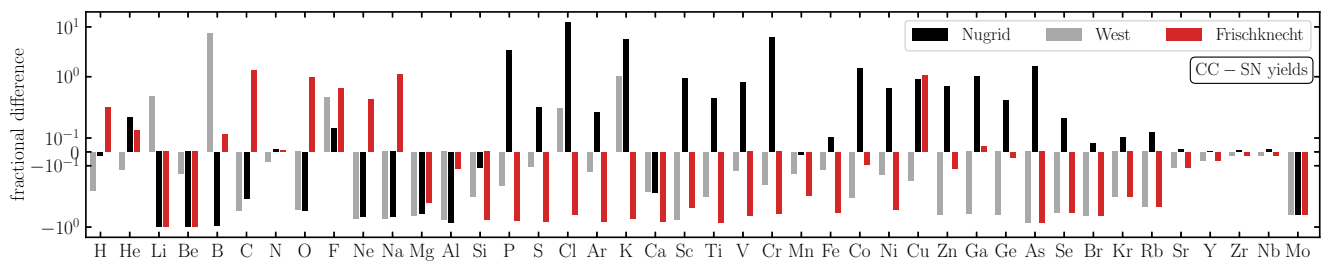


Figure D1. Same as Fig. 2 but this time comparing the CC SN yields from Frischknecht et al. (2016), West & Heger (in preparation) and from the Nugrid collaboration (Ritter et al. 2018b) to our fiducial yield-set.

This paper has been typeset from a \LaTeX file prepared by the author.