

Stellar Migration and Chemical Enrichment in the Milky Way

James W. Johnson,¹★ David H. Weinberg,^{1,2} Fiorenzo Vincenzo,² Jonathan C. Bird,³ Sarah R. Loebman,⁴ Alyson Brooks,⁵ Thomas R. Quinn,⁶ Charlotte R. Christensen,⁷ et al.?

¹ Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH, 43210, USA

² Center for Cosmology and Astroparticle Physics (CCAPP), The Ohio State University, 191 W. Woodruff Ave, Columbus, OH, 43210, USA

³ Department of Physics & Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN, 37235, USA

⁴ Department of Physics, University of California Merced, 5200 North Lake Rd., Merced, CA, 95343, USA

⁵ Department of Physics & Astronomy, Rutgers University, 136 Frelinghuysen Rd, Piscataway, NJ, 08854, USA

⁶ Department of Astronomy, University of Washington, Box 351580, Seattle, WA, 98195, USA

⁷ Department of Physics, Grinnell College, 1116 8th Ave., Grinnell, IA, 50112, USA

Accepted XXX; Received YYY; in original form ZZZ

ABSTRACT

We investigate the impact of stellar migration on galactic chemical evolution models, considering a handful of assumptions for the star formation history and time-dependence of radial migration based on a zoom-in, hydrodynamical simulation of galaxy evolution. We find that models for the time-dependence of radial migration impact the type Ia supernova rate as a function of both time and radius, inducing significant variability. We demonstrate that this is a means with which young, α -enhanced stars can arise naturally out of inside-out galaxy growth with stellar migration. We find that the observed age-[α /Fe] relation is well-fit by an inside-out star formation history, while the age-[O/H] and age-[Fe/H] relations are well-fit by a late starburst model; no one model investigated here fits both simultaneously. We also find that our models predict a dichotomy in [α /Fe], but that they do not resemble that observed in the Milky Way due to an overabundance of intermediate [α /Fe] stars. This suggests that inside-out galaxy growth combined with radial migration, even with a late starburst, is not conducive to forming the infamous bimodality as it is observed. We postulate that more dramatic evolutionary scenarios (e.g. a two-infall model) may be necessary to describe the observed results. In conducting this analysis, we developed and made use of newly released features in the **Versatile Integrator for Chemical Evolution** (VICE) which are built to handle these simulations under a wide variety of assumptions. VICE is publicly available at <https://pypi.org/project/vice>.

Key words: methods: numerical – galaxies: abundances, evolution, star formation, stellar content

1 INTRODUCTION

- Known for some time that stars undergo radial migration (e.g. Wielen et al. 1996), but to date there are only a handful of chemical evolution models that take this into account (e.g. Schönrich & Binney 2009; Minchev et al. 2013; Sharma et al. 2020). This effect can have a considerable impact on observed abundance trends by mixing stars that formed out of Galactic regions with significantly different enrichment histories, and if not taken into account, will impact the conclusions of chemical evolution studies.
- The Age-Metallicity Relation (AMR)

- Age-metallicity relation in the solar neighbourhood exhibits considerable intrinsic scatter (Edvardsson et al. 1993), usually

attributed to radial migration of metal-rich (metal-poor) stars formed at smaller (larger) Galactocentric radii (Sellwood & Binney 2002; Haywood 2008; Roškar et al. 2008b; Schönrich & Binney 2009).

– Feuillet et al. (2018) reveal that super-solar metallicity stars are statistically older than solar metallicity stars (their Fig. 3). Contrasts with simple one-zone models where enrichment proceeds alongside star formation yielding a monotonic AMR (e.g. Andrews et al. 2017; Weinberg et al. 2017). One-zone models of starbursts can produce non-monotonic AMR due to the effect of dilution (Johnson & Weinberg 2020), but they by construction do not predict multiple abundances at fixed age. Argued using the Weinberg et al. (2017) analytic models that this is a consequence of radial migration - for a smooth star formation history (SFH), the youngest stars at any given radius will have com-

* Contanct e-mail: johnson.7419@osu.edu

position reflective of the local equilibrium abundance, and only older stars will have had adequate time to migrate to a similar Galactocentric radius.

- The Young Alpha-Rich Population

- Population of young (~few Gyr), $[\alpha/\text{Fe}] \sim 0.1\text{-}0.2$ stars in the solar neighbourhood.

- Found using stellar ages estimated from carbon-to-nitrogen ratios (Martig et al. 2016), isochrone matching (Feuillet et al. 2018, 2019), and with the asteroseismic ages in the original APOKASC catalog (Chiappini et al. 2015; Silva Aguirre et al. 2018; Pinsonneault et al. 2014). Silva Aguirre et al. (2018) demonstrated that these stars have kinematics similar to the rest of the high- α population, and argued based on this that they may be the consequence of stellar mergers or mass transfer events, making truly old stars simply appear younger.

- Mor et al. (2019) infer a factor of ~ 2 enhancement in the SFH of the Milky Way ~ 2 Gyr ago by comparing population synthesis models to observed stellar luminosity functions and color-magnitude diagrams from Gaia data (Gaia Collaboration et al. 2018). Isern (2019) reach similar conclusions by modeling the white dwarf luminosity function in the solar neighbourhood with Gaia parallaxes. Motivated by these results, Johnson & Weinberg (2020) demonstrate using one-zone chemical evolution models that a recent starburst can produce young, α -enhanced stars. Caveat: burst would have had to be sufficiently localized such that the young, α -rich stars remain outliers from an otherwise monotonically decreasing age- $[\alpha/\text{Fe}]$ relation.

- Metallicity Distributions and the Abundance Gradient

- Known for some time that the inner regions of Milky Way-like spirals are more metal-rich than their outskirts (SDSS references: Nordström et al. 2004b; Daflon et al. 2009; Frinchaboy et al. 2013; Hayden et al. 2014). There are variations depending on what type of stars are considered and what metallicity tracers are used. This qualitative result is also seen in the gas-phase (see, e.g., the results from the CHAOS project using HII regions; Berg et al. 2015, 2020).

- We know from APOGEE data that the observed MDF has a metal-poor mode and is skew-positive in the outer galaxy (and conversely, metal-rich and skew-negative in the inner galaxy; Hayden et al. 2015). The variation with Galactocentric radius is usually attributed to radial migration of metal-poor (metal-rich) stars from larger (smaller) radii, the same argument that (Sellwood & Binney 2002) used to interpret the results of Edvardsson et al. (1993) on the observed AMR in the solar neighborhood.

- Abundance gradients have been the focus of many studies and are quantified rather extensively, but their origin is somewhat up for debate. Applying the analytic models of Weinberg et al. (2017), one could argue that the gradient arises out of variations in the local gas-phase equilibrium abundance as a function of Galactocentric radius. This could be due to variations in the efficiency of outflowing winds removing metals from the star forming reservoir due to a deeper gravitational potential in the inner regions of Milky Way-like spirals. Nidever et al. (2014) used this methodology to successfully reproduce the APOGEE abundance data. However, recent chemical evolution models have successfully replicated the gradient with no outflowing winds at all (e.g. Minchev et al. 2013; Spitoni et al. 2019). This is based on arguments from Melioli et al. (2008, 2009) and Spitoni et al. (2008, 2009), who studied Galactic fountains and found that ejected metals tend to reaccrete to a Galactocentric radius similar

to where they originated. With this result, some authors argue that such outflows do not significantly alter the chemical evolution of a Galactic disc. A notable difference between these two models is that the latter requires significantly lower nucleosynthetic yields to predict physically realistic abundances.

- The $[\alpha/\text{Fe}]$ bimodality

- Milky Way stars segregate themselves into the low- and high- α sequences, a bimodality found in, e.g., Gaia ESO (Recio-Blanco et al. 2014; Rojas-Arriagada et al. 2017), and APOGEE (Nidever et al. 2014; Hayden et al. 2015; Weinberg et al. 2019).

- Presence is well established, though origin a topic of intense debate.

- Notion that it could arise out of radial migration traces back to Schönrich & Binney (2009).

- Weinberg et al. (2017) models suggest increase in strength of the mass-loading factor η at late times would lower the equilibrium abundance, forming stars along the low- α sequence. This plus radial migration is yet unexplored.

- Spitoni et al. (2019) demonstrate that two-infall models can reproduce solar annulus data with good agreement. This plus radial migration is yet unexplored.

- Grand et al. (2018) find with Auriga (Grand et al. 2017) that early, accretion induced starburst populates the high- α sequence, followed by low-level, sustained star formation on low- α sequence, and a rapid transition between the two ensures chemical space relatively unpopulated. This would imply short τ_\star (see justification in Weinberg et al. 2017). Ongoing, low-metallicity gas accretion can also populate low- α sequence. Buck (2020) find results qualitatively similar to second scenario in NIHAO simulation suite (Wang et al. 2015; Buck et al. 2020). Hydrodynamical simulations take into account radial migration by construction.

- Clarke et al. (2019) show that star formation proceeded in clumps in an SPH simulation of an NFW halo using GASOLINE (Navarro et al. 1997; Wadsley et al. 2017). The clumps self-enrich, forming stars on the high- α sequence, while more spatially extended, smooth star formation populated the low- α sequence.

- No shortage of models that reproduce the dichotomy, yet only those done with hydrodynamical simulations and a handful of others have taken into account radial migration.

- This paper is meant to address the following simple question: When combining simple, conventional assumptions about the star formation and dynamical histories of the Milky Way, what observed results can be replicated?

2 METHODS

- To address the main goal of this paper, we make use of multi-zone chemical evolution models in which the Milky Way is described by a series of concentric annuli. Examples of similar methodologies already in the literature can be found in Matteucci & Francois (1989), Schönrich & Binney (2009), Minchev et al. (2013), and Sharma et al. (2020). Here we entertain a handful of models describing the star formation history and the time-dependence of radial migration, then simply assess their predictions in comparison to the results discussed in § 1.

- Multizone models are a means of extending traditional one-zone models to take into account the motions of stars and gas by allowing the exchange of both between zones.

- Our fiducial model has an inside-out star formation history (SFH) with parameters motivated by various observational constraints. Motivated by the findings of Mor et al. (2019) and Isern (2019), we construct a model with a late burst of star formation for comparison. A starburst of similar form was investigated in one-zone models in Johnson & Weinberg (2020). For comparison, we also construct a model in which the SFH is constant with time at fixed Galactocentric radius. This is interesting primarily from a theoretical perspective; it removes the effects of a time-varying SFH and quantifies what is predicted when there is ongoing star formation with radial migration.

- To include the effects of stellar migration, we make use of data from a hydrodynamical simulation, discussed in § 2.1. This allows us to adopt the dynamical history of a much more sophisticated simulation at face value, providing us with a model for radial migration that requires no free parameters.

- The Schönrich & Binney (2009), Minchev et al. (2013), and Sharma et al. (2020) studies chose a model for radial migration that was based on dynamical arguments, which introduced free parameters that then required fitting to data. An advantage of our decision to make use of a hydrodynamical simulation over dynamical arguments is that it's unclear the extent to which fitting to observed data biases the model into agreement with parts of the data that weren't involved in the fit. A hydrodynamical simulation, while still an approximation of reality, is motivated by first principles in that it models the gravitational interactions of star particles with gas and other star particles as well as the diffuse dark matter halo. Furthermore, we choose here a hydrodynamical simulation which is ran from cosmological initial conditions.

- The primary disadvantage of this choice is a consequence of the lack of free parameters - the model is rigid, and we cannot explore slight variations. However, in principle one could compare the predictions made by the same chemical evolution models but with different hydrodynamical simulations.

- To fulfill the goals of this paper, we develop and make use of newly released features within the **Versatile Integrator for Chemical Evolution** (VICE), which are designed to handle these kinds of simulation with a wide range of flexibility. We reserve description of VICE's algorithm for §2.2, first describing our sample of star particles from the hydrodynamical simulation.

2.1 The Hydrodynamical Simulation

- In this paper we make use of star particles from the h277 simulation (Christensen et al. 2012; Zolotov et al. 2012; Loebman et al. 2012, 2014; Brooks & Zolotov 2014). A synopsis of the detailed simulation parameters and cosmological model can be found in § 2 of Bird et al. (2020). We do not go into detail on that here, instead focusing on how we vet the sample of star particles for use in our chemical evolution models.

- We find that only ~3% of h277's star particles formed within the first 1 Gyr of its 13.7 Gyr run, indicating that $T = 1$ Gyr is a much better estimate of the onset of star formation and thus the chemical evolution in h277 than $T = 0$. We thus remove the star particles that were born in the first Gyr of evolution, and subtract 1 Gyr from their formation times, thus running our disc models for 12.7 Gyr. As a result, our simulations run from a redshift of $z \approx 3$ to the present day.

- Of particular importance to have accurately measured formation radii for each star particle that we use in our analysis. h277 did not record the exact birth radius of each star particle; however,

each star particle does have an accurate age at each snapshot. If a star particle is sufficiently young in the first snapshot it appears in, then its radius in that snapshot is a reasonable approximation for its birth radius, since its orbit is unlikely to have changed significantly in a short enough time interval. Therefore, we restrict our sample to those with an age at first snapshot of ≤ 150 Myr, and adopt their Galactocentric radius at first snapshot as the birth radius. Also conducted the analysis with age at first snapshot of ≤ 50 Myr, and found similar results, suggesting this choice does not impact our conclusions. In practice, we find that 150 Myr also provides us with an adequately large number of star particles to sample from.

- Based on a kinematic decomposition performed on the present-day phase space distribution of the h277 star particles, we exclude halo stars, but include bulge, pseudobulge, and thin- and thick-disc stars. Even though we are not modeling the abundances in the bulge in this paper, the bulge is believed to have formed in-situ (citation), while a significant fraction of halo stars are believed to have been accreted (citation). It's also likely that the Milky Way bulge and disc exchanged statistically significant amounts of stars and gas (citation), coupling the chemical evolution of the two at least slightly. This effect can be addressed in our models by simply including these star particles in our sample. This yields a sample of 3,019,521 star particles from h277.

- h277 had a transient bar during its evolution, but does not have a bar at $z = 0$. This could mean that the star particles in h277 migrated in a manner which does not reflect the dynamical history of the Milky Way, but at present inferring a star's birth radius observationally is notoriously difficult anyway (citation), so its likely the impact of this decision is within current observational uncertainties. Nonetheless, it is worth noting that the Minchev et al. (2013) model, and by extension the Minchev et al. (2014) and Minchev et al. (2017) studies as well, selected a hydrodynamical simulation specifically so that it would have a bar at $z = 0$. A detailed investigation of the impact of bar evolution on stellar migration and thus chemical evolution is outside the scope of this paper, though can be conducted in principle by simply swapping the birth times and initial and final radii of h277 within VICE, rerunning the simulations, and comparing the results.

- Distributions of final radii in bins of birth radii and age are shown in the top row of Fig. 1. Conversely, the bottom row shows distributions of birth radii in bins of final radii and age.

- Focusing on the top row of panels in Fig. 1, we note that for stars born at any radius and time, the distribution of final radii is still peaked near the birth radius. With increasing age, it appears the mode final radius may move slightly inward. The tails of the distributions to large radii are relatively age-independent - some differences between the 0 - 2 and 8 - 10 Gyr age bins, but not much. However, the tails of the distributions to small radii are not age-independent, and move toward smaller R_{gal} with increasing age. This suggests that radial migration inward and outward occur on different timescales, in particular that inward migration is slower than outward migration. By extension this may suggest that inward and outward migration are tied to different physical processes. Roškar et al. (2008a) demonstrate using a cosmological simulation that resonant scattering at corotation causes stars to move outward and gas to move inward. It's possible that radial migration inward has different origins.

- Focusing on the bottom row of panels in Fig. 1, we note that the oldest stars at any Galactocentric radius at the present day were overwhelmingly born at smaller radii. The youngest stars, however, were overwhelmingly born at comparable radii, and the

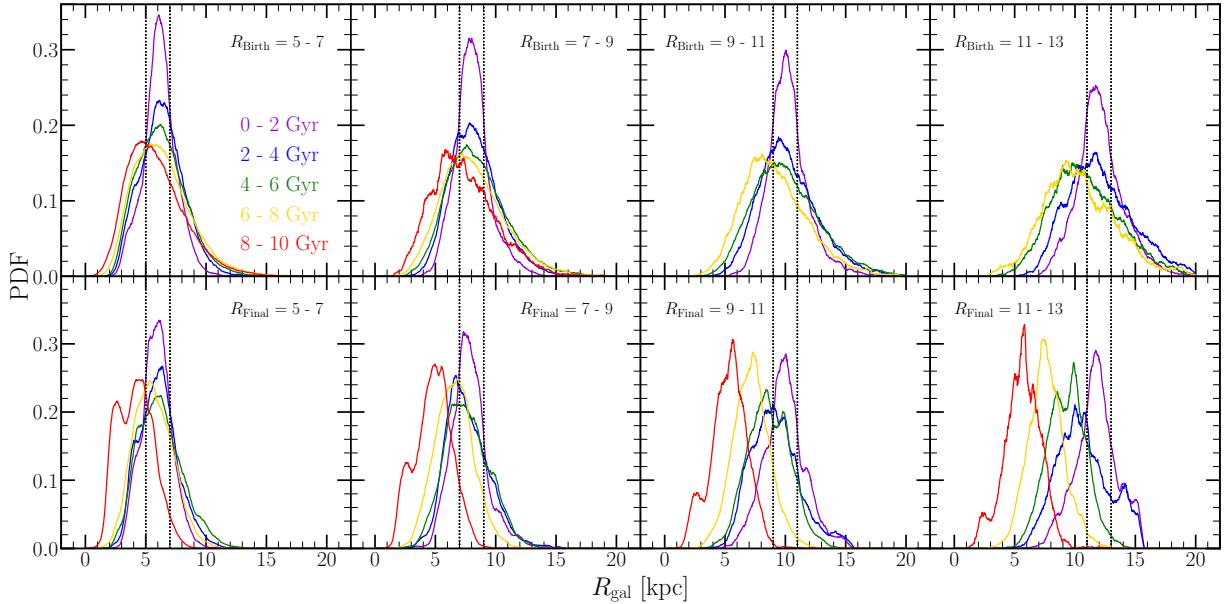


Figure 1. Radial distributions of our candidate analog star particles from h277. In the top row, we show distributions of *final* radius in bins of birth radius and age. In the bottom row, we show distributions of *birth* radius in bins of final radius and age. Each bin in Galactocentric radius is shown in its own panel, denoted in text at the top of each panel and by vertical black dashed lines. We color-code the distributions according to the age of the star particles, denoted by the legend in the upper left panel. We smooth all distributions with a box-car width of 0.5 kpc to improve clarity. We omit the distributions for 8 - 10 Gyr old stars born in the 9 - 11 and 11 - 13 kpc bins due to an insufficient number of star particles with which to calculate the distribution.

stars of intermediate ages simply span the range in radii between the two. With increasing radius, the differences in the mode of the birth radius distribution between age bins gets larger.

– The differences in distributions shown in the top and bottom panels boils down to stellar surface density being a strong function of Galactocentric radius. Take for example the age = 8 - 10 Gyr bin in both $R_{\text{Birth}} = 5 - 7$ kpc and $R_{\text{Final}} = 11 - 13$ kpc bins (i.e. the red curves in the top-left and bottom-right panels). For these old stars born at 5 - 7 kpc, 11 - 13 kpc is far down the tail of the R_{Final} distribution, and yet 5 - 7 kpc is the mode R_{Birth} of all old stars presently at these radii. This implies that even though the majority of 8 - 10 Gyr old stars with $R_{\text{Final}} = 11 - 13$ kpc were born at 5 - 7 kpc, they make up only a small fraction of the stars with similar birth radii. This is only possible if the stellar surface density gradient is sufficiently steep, which it is known to be (e.g. Bland-Hawthorn & Gerhard 2016).

• Fig. 1 also demonstrates that the numbers of stars that migrated inward and outward are comparable in h277. Taking a $|\Delta R_{\text{gal}}| \geq 500$ pc between birth and final radii as the criterion for migrating inward or outward, indeed we find that in our sample, 27% migrated inward, 29% migrated outward, and the remaining 44% stayed near their birth radius. These are global percentages.

– In all bins of birth radius, a good first-order estimate of the probability density that a star has a final radius in the same bin is ~ 0.3 . With bins in radius on this plot of 2 kpc, this suggests that $\sim 40\%$ of stars migrate significantly by the time they’re ~ 2 Gyr old. If the SN Ia DTD is a $t^{-1.1} \approx t^{-1}$ power-law as suggested by recent observations (e.g. Maoz & Mannucci 2012; Maoz & Graur 2017), then we expect similar numbers of SN Ia events to occur with delay times between 1 and 10 Gyr as we do between 100 Myr and 1 Gyr. With such an extended DTD, and the distributions in final radius shown in the bottom panel of Fig. 1, its

possible that white dwarfs migrate significant distances before producing a SN Ia event. Indeed, in the ASAS-SN bright supernova catalog, $\sim 10\%$ of supernovae are seen at > 10 kpc from their host galaxies (Holoi et al. 2019). While this catalog is for *all* supernovae, the majority of SN events are type Ia anyway. Based on this, it’s reasonable to expect that the migration of nucleosynthetic yields may proceed alongside stellar migration, an effect which is often neglected (e.g. Minchev et al. 2013, and the application of the Weinberg et al. 2017 analytic models in Feuillet et al. 2018). In this paper, we relax this assumption. The application of the h277 star particle data to our model for radial migration and the time dependence thereof is discussed in the next section.

2.2 Radial Migration

- To simulate our models, we develop and make use of VICE’s `milkyway` object, an extension of a more general object named `multizone`. The `multizone` object is at its core an array of `singlezone` objects, which are designed to handle simulations of one-zone models and were the focus of Johnson & Weinberg (2020), VICE’s initial release paper. The `multizone` object then affords the user the ability to move gas between zones with any given time dependence, and to assign all stellar populations any given zone at any given time following their birth, effectively allowing arbitrarily complex zone and migration schema. The `milkyway` object is a user-friendly extension of this which enforces a zone configuration in which the Galaxy is modeled as a series of concentric annuli of uniform width ΔR_{gal} . As defaults, it adopts the stellar migration model described in this section and our star formation law described in §2.6.

- Stars in VICE are stand-ins for entire stellar populations, and in the `milkyway` object, are said to be in a given zone/annulus if

their radius is between the inner and outer edges. At all times, their nucleosynthetic products and returned envelopes are placed in the ISM of the annulus that they are in *at that time*.

- Where hydrodynamical and N-body simulations of galaxy evolution often use star particles of a fixed mass, VICE forms a fixed number of stellar populations in a given zone at a given timestep, and allows their masses to vary to account for variations in the star formation rate. The mass of stars formed in a given zone is divided evenly among the stellar populations that form during a given timestep.

- Algorithm based on initial and final radii of star particles in hydrodynamical simulations, for which we've taken h277 as discussed in the previous section.

- In modeling the Milky Way as a series of concentric annuli, VICE's `milkyway` object assumes stellar populations to be born at the centres of each annulus. For a stellar population born at time T and Galactocentric radius R_{gal} , it searches for star particles from h277 that formed at $T \pm 300$ Myr and $R_{\text{gal}} \pm 250$ pc. It then randomly selects a star particle from this subsample with no bias to act as an *analog*. Adopts the final radius of the analog star particle at face value, and moves from its radius of birth to the final radius at $T = 12.7$ Gyr with some time dependence. If no candidate analogs are found, search is widened to $T \pm 600$ Myr and $R_{\text{gal}} \pm 500$ pc. If still no candidate analogs are found in this widened search, the stellar population is assumed to remain at its birth radius **and to have a final height above the disc midplane of $z = 100$ kpc**. These stellar populations are few in number. In our fiducial model with an inside-out SFH (see § 2.4), these stellar populations account for $\sim 4.9\%$ of the simulated stellar populations by number, but only $\sim 0.56\%$ by mass. Making up such a small portion of the mass budget, it's not likely they contribute significantly to the chemical evolution of our model Galaxy.

- We have zones for each annulus from $R_{\text{gal}} = 0$ to 20 kpc; there are no zones off the midplane. This retains the assumption that the gas-phase abundances are vertically and azimuthally well-mixed. Each stellar population simply assumes the final height $|z|$ above the disc midplane of its assigned analog, though this doesn't enter into the simulations at all; it's purely post-processing information.

- When a star particle is assigned as an analog to a stellar population in our simulations, it is *not* thrown out of the sample of candidate analogs, in theory allowing a star particle to act as an analog for multiple stellar populations. This is a difference between the Minchev et al. (2013) model and ours. Where we assign star particles to stellar populations for which VICE then calculates a stellar mass and composition based on an assumed star formation history, the Minchev et al. (2013) model assigns compositions to star particles.

- Our migration models

- All neglect radial gas flows, this paper instead focusing on four simple models for how the radius of a stellar population's orbit changes with time. Investigation of our models with radial gas flows would be an interesting extension to this paper.

- **Post-processing:** Stars stay where they are born until the final timestep. Based on the assumption that stellar populations do not contribute to enrichment beyond their birth radius (e.g. Minchev et al. 2013). Each annulus simulated as a one-zone model independent of all other zones. Shown in black dotted line in Fig. 2.

- **Sudden:** A random number is drawn from a uniform distribution between the time of birth and the present day. That time

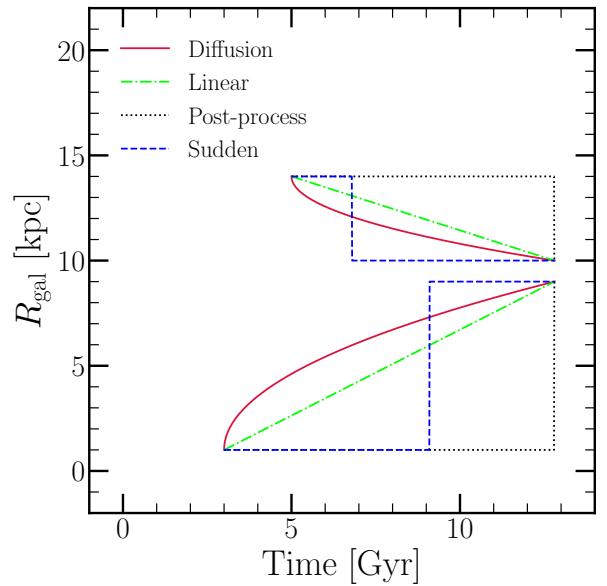


Figure 2. A diagram illustrating how Galactocentric radius changes with time for two stellar populations under our migration schema: diffusion (crimson, solid), linear (lime, dot-dashed), post-process (black, dotted), and sudden (blue, dashed). Here the initial and final radii and birth times are chosen at random for illustrative purposes. With the initial and final Galactocentric radii of a stellar population, its birth time, and one of these assumptions about the time-dependence of migration, the Galactocentric radius at all times is known.

is taken to be the time of instantaneous migration to the present-day annulus. Emulates a scenario in which a single dynamical interaction changes the orbital radius of a star. Can be thought of as a time-dependent extension of the post-processing scenario. Shown in blue dashed line in Fig. 2.

- **Diffusion:** A case in which stars migrate in a continuous, time-dependent manner. If stars move to their final radii via a random walk, then the mean displacement of stars that migrate similar distances would scale with $\sqrt{\text{age}}$. This is shown in the red solid line in Fig. 2. This is our fiducial migration model; we present results using this model except where otherwise noted. It is so named because it corresponds to a scenario in which the random walk carries out the diffusion of angular momentum.

- **Linear:** A simple variation of the diffusion model. Has no physical motivation other than numerical ease, but together with the diffusion model constitutes a test of how sensitive our models are to the assumed time-dependence of stellar migration. Demonstrated by green dot-dashed line in Fig. 2.

- Further details on the implementation of the `milkyway` object, the more general `multizone` object, and other features of VICE can be found in its documentation.¹

- We have a sample of 3,019,521 candidate analog star particles from h277, only $\sim 57\%$ of which are disc stars. Since we're modeling the thin and thick disc populations here, ~ 1.72 million is a much better estimate of the number of analogs that we can realistically sample from. We let star formation extend out to 15.5 kpc, beyond which we force the star formation rate to zero at all times. VICE does form stellar populations at all timesteps in the

¹ <https://vice-astro.readthedocs.io>

annuli; they simply have zero mass and thus do not contribute to the chemical evolution or recycling, though they do contribute to the computational overhead. We take $\Delta R_{\text{gal}} = 100$ pc as the width of each annulus from $R_{\text{gal}} = 0$ to 20 kpc and a timestep size of $\Delta T = 10$ Myr from $T = 0$ to 12.7 Gyr (see discussion in § 2.1 for our choice of time-interval). With the resulting 200 zones and 1,271 timesteps, we let VICE form $n = 8$ stellar populations per zone per timestep, resulting in 2,033,600 stellar populations with predicted masses and abundances, 1,574,800 of which form between $R_{\text{gal}} = 0$ and 15.5 kpc and have nonzero mass, reasonably within the limit of what we can sample. These simulations run in ~ 2 hours and take up ~ 235 MB of disc space per output, including the extra data that we record for each stellar population’s analog star particle.

2.3 Nucleosynthetic Yields, Outflows, and Recycling

- *Fractional net* yields, as required by VICE. Recycling is implemented separately, such that stellar envelopes are returned to the ISM at the birth composition of each stellar population. VICE takes the general approach of returning previously produced material according to a stellar lifetime model and injecting newly produced nucleosynthetic products as described below. Also clarify that these are not *effective* yields in that outflows are also implemented separately.

- CCSN events are assumed to occur immediately following the formation of their progenitor stars. This is an adequate approximation, because the lifetimes of massive stars are short compared to the relevant timescales for galaxy evolution. For the most massive stars, the lifetimes are comparable to the timestep size that we take use in these simulations anyway. This assumption implies a linear relationship between CCSN enrichment and the star formation rate:

$$\dot{M}_x^{\text{CC}} = y_x^{\text{CC}} \dot{M}_\star \quad (1)$$

- Physically, the CCSN yield y_x^{CC} is the fraction of a stellar population’s initial mass that is converted into some element x and ejected to the ISM, neglecting outflows.

- VICE includes functionality with which to calculate the value of y_x^{CC} for a handful of nucleosynthetic yield studies. Details can be found in § 2 of [Johnson & Weinberg \(2020\)](#) and in VICE’s documentation. A detailed investigation of CCSN nucleosynthetic yields which releases new features to these calculations will be presented in Griffith et al. (2021, in prep).

- Take $y_O^{\text{CC}} = 0.015$ and $y_{\text{Fe}}^{\text{CC}} = 0.0012$ from [Johnson & Weinberg \(2020\)](#), who in turn adopt these values from [Weinberg et al. \(2017\)](#).

- SN Ia products injected with a $t^{-1.1}$ delay-time distribution (DTD) with a minimum delay-time of $t_D = 150$ Myr. This is the default DTD in VICE, adopted in [Johnson & Weinberg \(2020\)](#), and is suggested by recent observational results comparing the cosmic SN Ia rate to the cosmic SFH ([Maoz & Mannucci 2012](#); [Maoz & Graur 2017](#)). In a one-zone model at times $t > t_D$, the enrichment rate for some element x can be expressed as the product of some yield y_x^{Ia} and the star formation history weighted by the DTD:

$$\dot{M}_x^{\text{Fe}} = y_x^{\text{Fe}} \langle \dot{M}_\star \rangle_{\text{Ia}} \quad (2a)$$

$$= y_x^{\text{Fe}} \frac{\int_0^{t-t_D} \dot{M}_\star(t') R_{\text{Ia}}(t-t') dt'}{\int_{t_D}^{t_{\max}} R_{\text{Ia}}(t-t') dt'} \quad (2b)$$

where R_{Ia} is the DTD itself, in units of $M_\odot^{-1} \text{Gyr}^{-1}$. Like the CCSN yield, y_x^{Ia} is simply the fraction of a single stellar population’s mass that is converted into the element x over the ensuing SN Ia duty cycle. It can also be expressed as an integral over the DTD:

$$y_x^{\text{Ia}} = m_x^{\text{Ia}} \int_{t_D}^{t_{\max}} R_{\text{Ia}}(t) dt = m_x^{\text{Ia}} \frac{N_{\text{Ia}}}{M_\star} \quad (3)$$

where the m_x^{Ia} is the mass of some element x produced in a single SN Ia event on average, and the integral evaluates to the mean number of Ia events N_{Ia} per mass of stars formed M_\star .

- VICE enforces $t_{\max} = 15$ Gyr always, though provided one is consistent with equations (3) and (2b), the results are independent of t_{\max} since the integrals cancel.

- Extending this to multi-zone models is simple - rather than an integral over the star formation history of a given annulus, the rate becomes a summation over all stellar populations that are in a given zone at some time:

$$\dot{M}_x^{\text{Ia}} = y_x^{\text{Ia}} \frac{\sum_i M_i R_{\text{Ia}}(\tau_i)}{\int_{t_D}^{t_{\max}} R_{\text{Ia}}(t) dt} \quad (4)$$

where M_i and τ_i are the mass and age of the i ’th stellar population, respectively.

- Initially adopt $y_O^{\text{Ia}} = 0$ and $y_{\text{Fe}}^{\text{Ia}} = 0.0017$ from [Johnson & Weinberg \(2020\)](#) and [Weinberg et al. \(2017\)](#). However, in practice we find that the e-folding timescales of star formation in our models are sufficiently long (see discussion in § 2.4) such that our fiducial, inside-out SFH model predicts $[\text{O}/\text{Fe}] \approx +0.05$ for young stars. We therefore multiply this value by a factor of $10^{0.1}$, adopting instead $y_{\text{Fe}}^{\text{Ia}} = 0.00214$ so that our fiducial model predicts a late-time $[\text{O}/\text{Fe}]$ ratio in better agreement with observations. This change is likely within the uncertainties in nucleosynthetic yields anyway.

- Our model assumes that all supernova yields of O and Fe are independent of metallicity. While this appears to be empirically true for CCSN yields of these elements (e.g. [Weinberg et al. 2019](#); [Griffith et al. 2020](#)), recent evidence by [Brown et al. \(2019\)](#) suggests that the local specific SN Ia rate shows a strong, inverse dependence on galaxy stellar mass. They argue that this may imply a metallicity dependent R_{Ia} that produces more SN Ia events at early times when the metallicity is low. While we adopt a metallicity independent $y_{\text{Fe}}^{\text{Ia}}$ in this paper, VICE allows users to specify any functional form for $y_{\text{Fe}}^{\text{Ia}}$ as a function of the total abundance by mass Z , allowing future studies of the impact of this to proceed straight-forwardly.

- Similar to SN Ia, the equation detailing asymptotic giant branch (AGB) star enrichment changes from an integral over the SFH to a summation over the stellar populations in a given zone at some time. VICE in its current iteration forces AGB enrichment to be included in its simulations, though the yields of O and Fe from AGB stars are negligible compared to their supernova yields, and the results presented in this paper can be interpreted as if they were zero.

- The algorithm is based on determining the fraction of a single stellar population’s mass that is in the form of stars that will have ended their post-main-sequence lifetime during the next timestep. The mass of those stars and their birth metallicity can then be used to determine a yield given a table of y_x^{AGB} values sampled on a grid of stellar mass and metallicity, for which we

take values from the FRUITY database of net yields (Cristallo et al. 2011). Since these yields are not important for this paper, however, we do not go into further detail. More information on VICE’s treatment of AGB enrichment can be found in its documentation and in § 2.6 of Johnson & Weinberg (2020).

- Outflows are directly tied to nucleosynthetic yields, because the absolute scaling of nucleosynthetic yields and the efficiency with which galactic winds remove them from the star forming reservoir determine the *effective* yield: the amount of metals that are not only produced but added to the star forming reservoir.

– Based on studies of galactic fountains demonstrating that ejected metals tend to fall back to a similar radius at which they were produced (e.g. Melioli et al. 2008, 2009; Spitoni et al. 2008, 2009), a number of chemical evolution models (e.g. Minchev et al. 2013; Spitoni et al. 2019) neglect outflows, arguing that they are not a necessary ingredient to chemical evolution models. However, with the yields we’ve adopted, our models require significant metal ejection to predict realistic abundances. For this reason, studies which neglect outflows adopt significantly lower nucleosynthetic yields as a consequence.

– There is a near one-to-one degeneracy between the scaling of nucleosynthetic yields and the efficiency of metal ejection in galaxies. With nucleosynthetic yields that were conventional at the time, Dalcanton (2007) demonstrated that significant outflows are required to lower effective yields and predict abundances in agreement with observations. This is a solution to the so-called “G-dwarf problem” in which closed-box models of chemical evolution (see the review in, e.g., Tinsley 1980) catastrophically over-predict the number of super-solar metallicity stars (van den Bergh 1962; Pagel & Patchett 1975). However, recent findings with regard to black hole formation and stellar explodability have demonstrated that many massive stars simply do not produce a supernova event (see theoretical discussion by, e.g., Pejcha & Thompson 2015; Sukhbold et al. 2016, and observation evidence from Gerke et al. 2015; Adams et al. 2017; Basinger et al. 2020). This challenges the results of Dalcanton (2007), suggesting that nucleosynthetic yields are intrinsically low, and the extent to which theoretically predicted stellar explodability models reduce the need for outflowing winds remains an unanswered question for future studies.

- Mass-loading factor η defined as the ratio of mass outflow rates from the ISM to the SFR:

$$\eta \equiv \frac{\dot{M}_{\text{out}}}{\dot{M}_{\star}} \quad (5)$$

Weinberg et al. (2017) demonstrate that, to first order, this parameter and the yields of a given element set the late-time equilibrium abundance. Here we adopt a scaling of η with Galactocentric radius R_{gal} such that the equilibrium abundance as a function of R_{gal} reproduces a metallicity gradient in agreement with observations.

- A fundamental observable, the observed metallicity gradient in the Milky Way has been the focus of a considerable number of studies to date.

– Nordström et al. (2004b) find a gradient of -0.099 kpc^{-1} in [Fe/H] in main sequence stars from GCS (Nordström et al. 2004a; Holmberg et al. 2007).
 – Daflon et al. (2009) find a gradient of -0.04 kpc^{-1} in [S/H] in OB stars.
 – Frinchaboy et al. (2013) find a gradient of -0.09 kpc^{-1} in [M/H] in open clusters.

– Hayden et al. (2014) also find -0.09 kpc^{-1} in [M/H] for $R_{\text{gal}} \gtrsim 6 \text{ kpc}$ for low- α stars. For $R_{\text{gal}} \lesssim 6 \text{ kpc}$, they find a relatively flat gradient.

– Weinberg et al. (2019) finds -0.06 kpc^{-1} in mode([Mg/H]) for upper red giant branch disc stars (see their Fig. 23).

- The procedure outlined in this section makes two assumptions: 1) that the mode abundance at a given radius reflects the equilibrium value at that radius, and 2) that radial migration does not significantly impact the overall form of the gradient. We demonstrate that this holds in our simulations in § 4.

- For α elements, Weinberg et al. (2017) defines the equilibrium abundance under a constant SFH as:

$$Z_{\alpha,\text{eq}} = \frac{y_{\alpha}^{\text{CC}}}{1 + \eta(R_{\text{gal}}) - r} \quad (6)$$

where r is the recycling parameter (≈ 0.4 for the sake of this scaling with a Kroupa IMF; see discussion near the end of this section). Solving for $\eta(R_{\text{gal}})$ yields:

$$\eta(R_{\text{gal}}) = \frac{y_{\alpha}^{\text{CC}}}{Z_{\alpha,\text{eq}}} + r - 1 = \frac{y_{\alpha}^{\text{CC}}}{Z_{\alpha,\odot}} 10^{-\text{mode}([\text{X}/\text{H}](R_{\text{gal}}))} + r - 1 \quad (7)$$

- We assume a slope of -0.08 kpc^{-1} , in tentative agreement with the previous studies that quote the slope of the gradient mentioned above. To set the normalization, we assume the mode([X/H]) to be ~ 0.3 at $R_{\text{gal}} = 4 \text{ kpc}$, since this would produce mode([X/H]) ≈ 0 at $R_{\text{gal}} \approx 7 - 9 \text{ kpc}$. This results in the following form for η as a function of Galactocentric radius:

$$\eta(R_{\text{gal}}) = \frac{y_{\alpha}^{\text{CC}}}{Z_{\alpha,\odot}} 10^{(-0.08 \text{ kpc}^{-1})(R_{\text{gal}} - 4 \text{ kpc}) + 0.3} + r - 1 \quad (8)$$

where we adopt our yield of O for y_{α}^{CC} and the solar abundance of O of $Z_{\odot,\odot} = 0.00572$ from Asplund et al. (2009).

- This does assume a uniformly linear gradient at all R_{gal} . Hayden et al. (2014) do find the gradient in [M/H] to flatten for $R_{\text{gal}} \lesssim 6 \text{ kpc}$, challenging this assumption. However, this procedure can be easily repeated for any desired gradient, since the functional form simply goes into the power of 10 in equation (8).

- The top panel of Fig. 3 plots this scaling of η with R_{gal} , highlighting a value of ~ 2.15 for the solar circle in the red dotted line.

- Weinberg et al. (2017) defined the *cumulative return fraction* r to be the fraction of a single stellar population’s initial mass that has been returned to the ISM. In detail, r is a complicated function of the age of a stellar population τ which depends on the assumptions about stellar lifetimes, the IMF, and an initial-remnant mass relation (e.g. Kalirai et al. 2008).

– Weinberg et al. (2017) demonstrated that the assumption of *instantaneous recycling* is a reasonable approximation, where instantaneous recycling refers only to *previously produced* nucleosynthetic products. We clarify that previous studies have made this approximation for newly produced material as well; while that assumption is valid for prompt sources like CCSNe, it fails for enrichment channels with significant DTDs, such as SNe Ia or AGB stars. Weinberg et al. (2017) demonstrate that $r = 0.4$ (0.2) is a good approximation for a Kroupa (Salpeter) IMF. However, this was demonstrated for smooth star formation histories and in parameter spaces conducive to analytic solution. Because numerical integration of $r(\tau)$ is neither challenging nor time-consuming, VICE by default implements continuous recycling by solving $r(\tau)$ at the beginning of each simulation given the IMF,

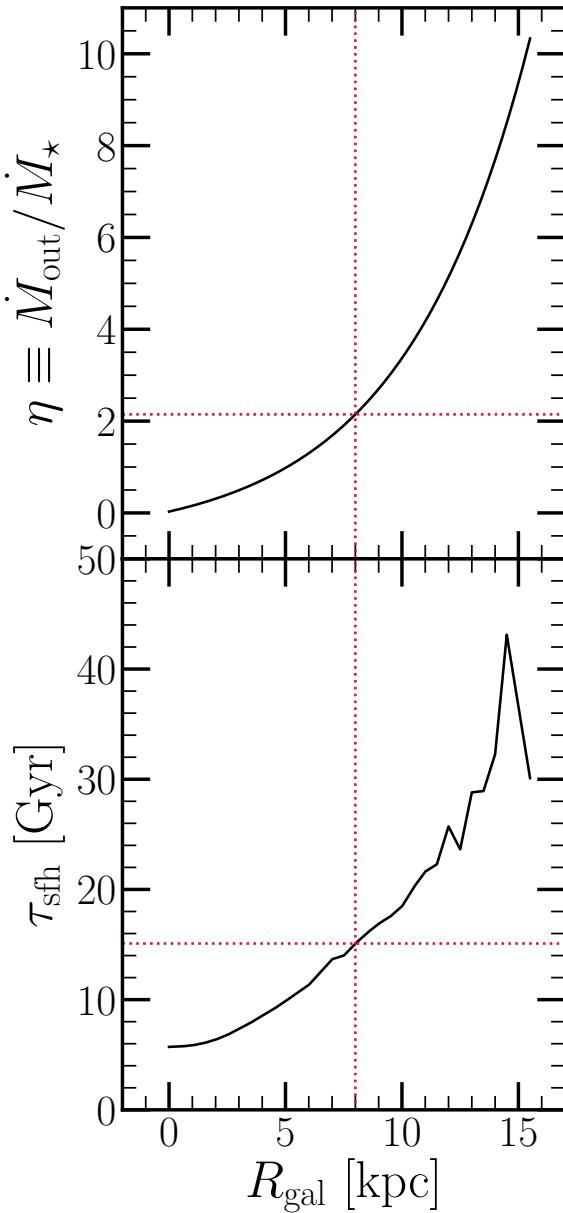


Figure 3. Top: Our implemented scaling of the mass loading factor η with Galactocentric radius (black) as defined by equation (8). **Bottom:** The e-folding timescales of the star formation histories of our model galaxies (black). These values come from a fit to the Σ_\star -age relation in bins of R/R_e for $10^{10.5} - 10^{11} M_\odot$ Sa/Sb Hubble type spiral galaxies as reported by Sánchez (2020, see discussion in § 2.4). The horizontal and vertical red dashed lines in both panels highlight a mass loading factor of $\eta \approx 2.15$ and a star formation timescale of $\tau_{\text{sfh}} \approx 15$ Gyr at an assumed radius of the sun of $R_\odot = 8$ kpc.

the Kalirai et al. (2008) initial-remnant mass model, and the assumption that $\tau = \tau_\odot(M/M_\odot)^{-3.5}$. We also assume a post-main sequence lifetime of 10% of the main sequence lifetime for all stars. Further details on the detailed form of $r(\tau)$ and its numerical integration within VICE can be found in its documentation.

- In a one-zone model, the recycling rate can be expressed as an

integral over the star formation history:

$$\dot{M}_x^r = \int_0^t Z_x(t') \dot{M}_\star(t') \dot{r}(t-t') dt' \quad (9)$$

where Z_x is the abundance by mass of the element x in the ISM at a time t' . Like SN Ia enrichment rates, extending this to multi-zone models changes the integral over the star formation history to a summation over the stellar populations in a given zone at some time:

$$\dot{M}_x^r = \sum_i Z_{x,i} M_i \dot{r}(\tau_i) \quad (10)$$

In the case of the gas recycling rate, the equation has the same form, but without the weighting by metallicity.

- We implement continuous recycling in our simulations, but make use of the instantaneous approximation in a handful of places in this paper (e.g. in our scaling of η with R_{gal} , and in approximating SN Ia rates in § 3).

2.4 Star Formation Histories

- VICE’s simulations runs in either infall, star formation, or gas mode, referring simply to which component of the evolutionary history the user has specified. The fiducial starburst models presented in Johnson & Weinberg (2020) ran in infall mode, but here we run things in star formation mode, since we are after specific forms for the star formation histories of our models.

• Appendix A presents justification of how we normalize the parameters of our star formation histories to produce a realistic model Galaxy at the end of the simulation. In short, it takes in a unitless description of the time-dependence of the SFH at a given Galactocentric radius, denoted $f(t|R_{\text{gal}})$, and a unitless description of the radial dependence of the present-day stellar surface density, denoted $g(R_{\text{gal}})$. In short, we integrate $f(t|R_{\text{gal}})$ with time for each annulus, assuming R_{gal} to correspond to the centre of the zone, and attach a prefactor to $f(t|R_{\text{gal}})$ at each R_{gal} such that the desired gradient is achieved with a total stellar mass similar to that of the Milky Way. This procedure neglects the impact of radial migration, assuming that it does not significantly alter the form of $g(R_{\text{gal}})$. We demonstrate that these assumptions hold in § 2.5, in which we also detail our adopted form of $g(R_{\text{gal}})$. As long as this assumption is not violated, the equation derived in Appendix A can be used to normalize the parameters of future models of spiral galaxies.

- We present four fiducial SFHs, which we dub “constant”, “inside-out”, “late-burst”, and “outer-burst”. They’re defined as follows:

- **Constant:** The SFH at a given radius is time-independent.

$$f_C(t|R_{\text{gal}}) = 1 \quad (11)$$

This case is of particular theoretical interest because it quantifies the effect of ongoing with star formation with radial migration, and no additional effects.

- **Inside-Out:**

$$f_{IO}(t|R_{\text{gal}}) = (1 - e^{-t/\tau_{\text{rise}}}) e^{-t/\tau_{\text{sfh}}} \quad (12)$$

We adopt this scenario over the traditional linear-times-exponential form of $f(t|R_{\text{gal}}) \sim t e^{-t/\tau_{\text{sfh}}}$, because the latter does not offer control over the position of the maximum. The form we adopt has a maximum near τ_{rise} , for which we adopt a value 2 Gyr everywhere in this paper. We find that this produces a peak in star formation at lookback times of ~ 10 Gyr, corresponding to a redshift of $z \approx 2$, which is around cosmic high noon. In

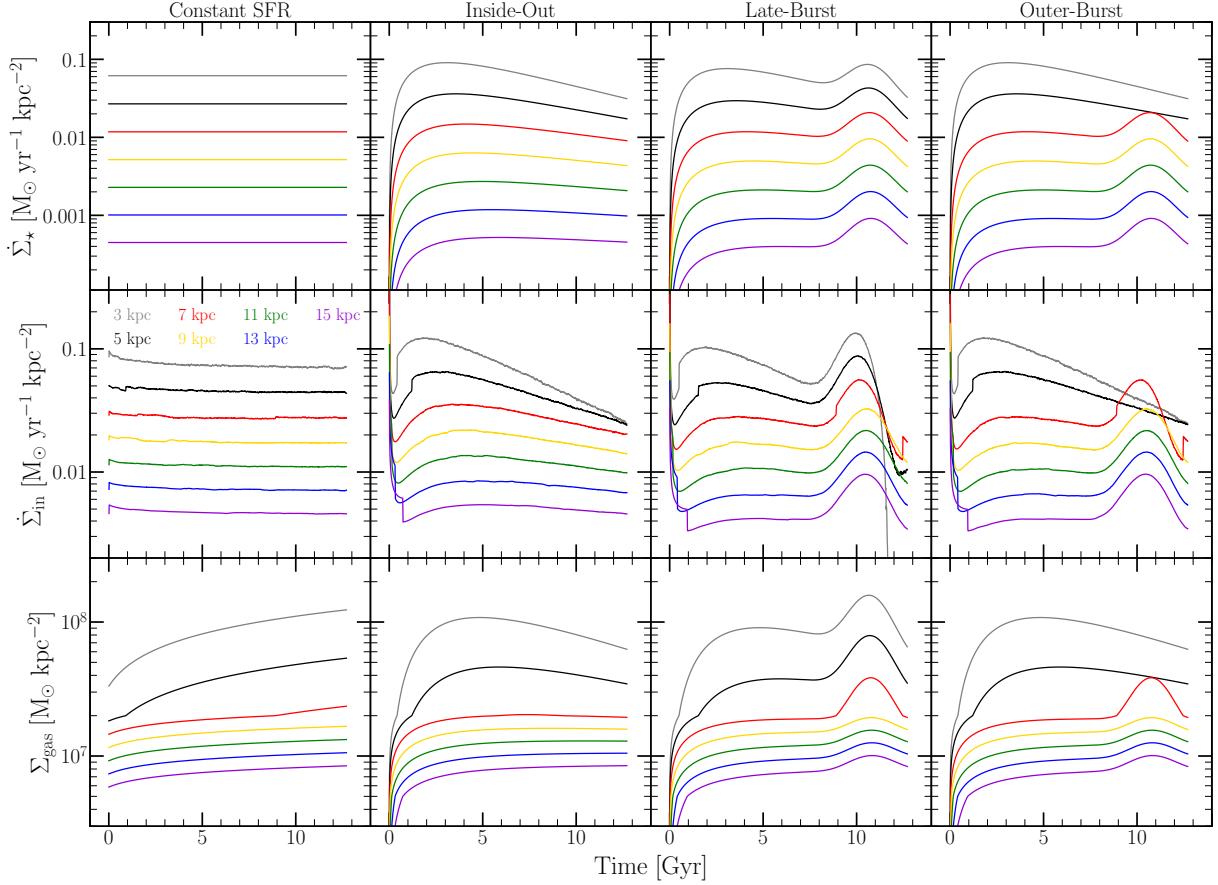


Figure 4. The surface densities of star formation $\dot{\Sigma}_*$ (top row), infall $\dot{\Sigma}_{\text{in}}$ (middle row), and gas Σ_{gas} (bottom row) as functions of simulation time for our four fiducial SFHs: constant (far left), inside-out (left middle), late-burst (right-middle), and outer-burst (far right). We plot curves for the annuli whose inner radii are 3 kpc (grey), 5 kpc (black), 7 kpc (red), 9 kpc (yellow), 11 kpc (green), and 13 kpc (purple) (see equations (11), (12), (13), and (14) for the mathematical definition of each SFH).

this paper τ_{sfh} is a function of R_{gal} here, and we discuss it briefly in a few paragraphs.

– **Late-Burst:** An inside-out SFH with a gaussian describing a starburst.

$$f_{\text{LB}}(t|R_{\text{gal}}) = f_{\text{IO}}(t|R_{\text{gal}})(1 + A_b e^{-(t-t_b)^2/2\sigma_b^2}) \quad (13)$$

A_b is a dimensionless parameter describing the strength of the starburst, t_b is the time of the local maximum in the SFH during the burst, and σ_b is the width of the gaussian describing it. Loosely motivated by the findings of [Isern \(2019\)](#) and [Mor et al. \(2019\)](#). Here we take $A_b = 1.5$, $t_b = 10.8$ Gyr, and $\sigma_b = 1$ Gyr.

– **Outer-Burst:** A variation of the late-burst model in which only $R_{\text{gal}} \geq 6$ kpc experience the starburst. Loosely motivated by findings of [Vincenzo & Kobayashi \(2020\)](#) where a hydrodynamical simulation of a Milky Way-like galaxy showed radially dependent infall.

$$f_{\text{OB}}(t|R_{\text{gal}}) = \begin{cases} f_{\text{IO}}(t|R_{\text{gal}}) & (R_{\text{gal}} < 6 \text{ kpc}) \\ f_{\text{LB}}(t|R_{\text{gal}}) & (R_{\text{gal}} \geq 6 \text{ kpc}) \end{cases} \quad (14)$$

- Derive e-folding timescales of star formation τ_{sfh} from the data in [Sánchez \(2020\)](#).

– They present the stellar surface density Σ_* as a function of age in bins of R/R_e for MaNGA galaxies, where R_e is the half-light radius. They conduct this analysis in bins of stellar mass and

for different Hubble types. Here take their $M_\star = 10^{10.5} - 10^{11} M_\odot$ bin for Sa/Sb spirals (i.e. Milky Way-like galaxies).

– Their reported Σ_* -age relation is not robust enough to get individual SFHs directly, but does allow the parameters of some fiducial mathematical model to be fit to the population-averaged trends. Assuming the $f_{\text{IO}}(t|R_{\text{gal}})$ form, we simultaneously fit the normalization and the e-folding timescale τ_{sfh} to these data. Although the normalization is irrelevant to our simulations and determined via the method outlined in Appendix A, we adopt the resulting $\tau_{\text{sfh}}-R_{\text{gal}}$ relation in our models. Results are shown in Fig. 3.

– Note that the star formation timescales are long, even for the solar annulus ($\tau_{\text{sfh}} \approx 15$ Gyr at $R_\odot = 8$ kpc) and especially for the outer galaxy. Beyond the solar annulus, SFHs are nearly constant after the initial rise at early times.

– With our adopted gradient (see §2.5, we know the present-day half-mass radius to be very near 4 kpc (this is just doing a couple integrals over the area of the disc). The findings of [García-Benito et al. \(2017\)](#) and [González Delgado et al. \(2014\)](#) suggest that half-light radii are marginally larger than half-mass radii. Based on equations (4) of [González Delgado et al. \(2014\)](#) relation the two for circular apertures, we expect our model galaxy to have a half-light radius near 5 kpc. We therefore adopt $R_e = 5$ kpc in calculating τ_{sfh} as a function of radius. As I understand it,

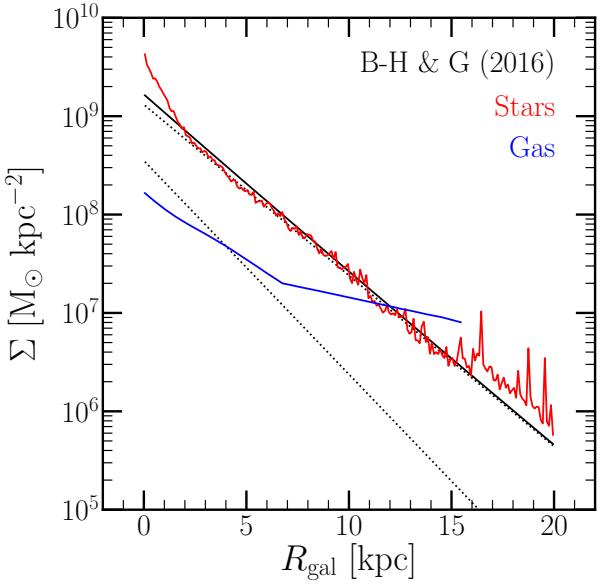


Figure 5. The surface density of gas (blue) and stars (red) as predicted by our inside-out SFH model. The dotted black line with the higher normalization denotes a thin disc profile with scale length $R_t = 2.5$ kpc; the other denotes a thick disc profile with scale length $R_T = 2.0$ kpc, and a ratio of $\Sigma_T / \Sigma_t = 0.27$ at $R_{\text{gal}} = 0$. The black line denotes the sum of the two; this is the surface density gradient of the Milky Way as reported by [Bland-Hawthorn & Gerhard \(2016\)](#), renormalized to imply the same total stellar mass within the disc.

there are some results that suggest this value for the Milky Way as well?

- Given the assumed star formation histories, the gas supply at all times is known via the SFE timescale τ_\star (discussed in § 2.6). With the amount of gas lost to star formation and outflows at each timestep, the infall rate is also known at all timesteps. The results of all three are shown in Fig. 4.

2.5 Surface Density Gradient

- As discussed in § 2.4, Appendix A presents justification of a recipe in which we select a unitless, unnormalized form describing the scaling of the stellar surface density with Galactocentric radius, denoted $g(R_{\text{gal}})$. Here we adopt the following double exponential form, describing the thin and thick discs of the Milky Way:

$$g(R_{\text{gal}}) = e^{-R/R_t} + \frac{\Sigma_T}{\Sigma_t} e^{-R/R_T} \quad (15)$$

where R_t and R_T are the scale radii of the thin and thick discs, respectively, and Σ_T / Σ_t is the ratio of their surface densities at $R_{\text{gal}} = 0$. In this paper, we adopt $R_t = 2.5$ kpc, $R_T = 2.0$ kpc, and $\Sigma_T / \Sigma_t = 0.27$ from the findings of [Bland-Hawthorn & Gerhard \(2016\)](#).

- Adopt the [Licquia & Newman \(2015\)](#) total stellar mass of $(6.08 \pm 1.14) \times 10^{10} M_\odot$. We’re including bulge star particles in our sample of candidate analogs from h277, so we model the entire stellar mass as opposed to just the disc, for which [Licquia & Newman \(2015\)](#) found $(5.17 \pm 1.11) \times 10^{10} M_\odot$.

- Surface density gradient from fiducial model with inside-out SFH, plotted in Fig. 5 (stars in red and gas in blue). Radial migration indeed didn’t change the overall scaling of Σ_\star with R_{gal} at the radii that we care about in this paper (3 kpc $\lesssim R_{\text{gal}} \lesssim 15$ kpc), only

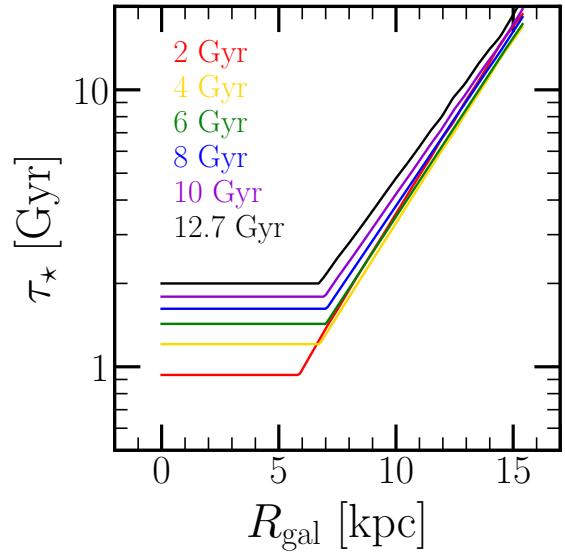


Figure 6. The star formation efficiency timescale τ_\star as a function of Galactocentric radius at simulation times of 2 Gyr (red), 4 Gyr (yellow), 6 Gyr (green), 8 Gyr (blue), 10 Gyr (purple), and 12.7 Gyr (i.e. the present day, black) predicted by our inside-out SFH model.

introducing scatter. Gas disc appears to be flatten at $R_{\text{gal}} \gtrsim 6$ kpc; this is a consequence of our model for star formation efficiency (see discussion in § 2.6). Black dotted lines show the individual thin and thick disc components that we adopt from [Bland-Hawthorn & Gerhard \(2016\)](#), with the solid line denoting the sum of the two, all re-normalized such that the integral over the surface area of the model Galaxy implies a total stellar mass in agreement with our adopted value from [Licquia & Newman \(2015\)](#).

2.6 Star Formation Efficiency

• The term “star formation efficiency” (SFE) is an overloaded term in the literature. In the star formation/feedback literature, it usually refers to the fraction of a molecular cloud’s mass which will eventually be converted into stars. In the chemical evolution literature, however, it usually refers to the inverse timescale relating the star formation rate within some gas reservoir and the mass of the gas reservoir itself: $\tau_\star \equiv \Sigma_{\text{gas}} / \dot{\Sigma}_\star$. High (Low) values of τ_\star indicate slow (fast) star formation and thus low (high) SFE; when we refer to SFE here, we’re referring to the definition based on this timescale.

- Potentially note that this is often referred to as the “depletion time” in star formation literature, and the [Weinberg et al. \(2017\)](#) definition of depletion time: $\tau_{\text{dep}} \equiv \tau_\star / (1 + \eta - r)$.
- This is a topic of intense observational inquiry, results of which we review in this section in the interest of obtaining an observationally motivated star formation law.

– In practice, many chemical evolution studies adopt a simple power-law for the Kennicutt-Schmidt law relating Σ_{gas} and $\dot{\Sigma}_\star$: $\dot{\Sigma}_\star \sim \Sigma_{\text{gas}}^N$, where [Kennicutt \(1998\)](#) infer $N = 1.4 \pm 0.15$ across a sample of quiescent spiral galaxies and infrared and circumnuclear starbursts. [de los Reyes & Kennicutt \(2019\)](#) derive a similar value of $N = 1.41 \pm 0.07$ for local quiescent spiral, dwarf, and low surface brightness galaxies, further demonstrating that much of the observed scatter is intrinsic.

– Liu et al. (2015) demonstrate that the derived value of N is sensitive to the assumed form of the CO-to-H₂ conversion factor (α_{CO}); for fixed α_{CO} , they find $N = 1.14 \pm 0.02$, but $N = 1.60 \pm 0.03$ for a continuously varying α_{CO} for a sample of 181 local galaxies with infrared luminosities spanning nearly five orders of magnitude. Much of the uncertainty in the detailed form of the star formation law is likely a consequence of the ongoing debate about α_{CO} (and other mathematical definitions of the conversion factor as well, such as X(CO); see discussion in Kennicutt & Evans 2012).

– In extending the de los Reyes & Kennicutt (2019) study to the most extreme infrared-luminous and circumnuclear starburst galaxies, Kennicutt & de los Reyes (2020) recently found a value of $N = 1.50 \pm 0.05$. Although they determine that a single power law is a reasonable recipe for models of Galaxy evolution, they demonstrate that there is evidence for significant breaks in both the power-law indeces and the zeropoints. They argue that these breaks are physical in origin, and speculate that it is tied to the transition from a mixed atomic and molecular ISM in normal disc galaxies to a nearly fully molecular ISM in starburst galaxies. A transition to a linear $\dot{\Sigma}_\star - \Sigma_{\text{gas}}$ relation at a surface density above which the molecular fraction $f_{\text{mol}} \approx 1$ is consistent with the finding that star formation proceeds only in the molecular phase (e.g. Leroy et al. 2008, 2013).

– Consistent with the finding that the intrinsic scatter in the observed $\dot{\Sigma}_\star - \Sigma_{\text{gas}}$ relation is physical in origin, Ellison et al. (2020a) demonstrate that this relation (among other fundamental correlations) shows significant galaxy-to-galaxy variations in both slope and normalization, suggesting that it is not universal between galaxies. Leroy et al. (2013) also find evidence that τ_{mol} , the value of τ_\star for a gas reservoir with $f_{\text{mol}} = 1$, is multivalued at fixed Σ_{H_2} , suggesting that the rate of star formation is set not only by the density of the gas reservoir, but other environmental parameters as well.

- Although we acknowledge the collective result that an individual galaxy would not obey the observed, population-averaged $\dot{\Sigma}_\star - \Sigma_{\text{gas}}$ relation, we decide to adopt such a prescription for our models here. Our motivation for doing so is two-fold.

– In practice, we find that our mathematical formalism for τ_\star does not significantly impact our conclusions; we demonstrate this in Appendix B. For this reason, a simple formalism is its own reward for our purposes.

– As discussed in § 2.4, we are running VICE in star formation mode; this simply means that we have specified the *star formation history* of the Galaxy, rather than the infall history or the ISM surface density as functions of simulation time and radius. While many galaxy evolution models use these observed star formation laws to infer the star formation rate from the ISM surface density, here we are doing the inverse - inferring the ISM surface density from the star formation rate. We therefore do not know the properties of the ISM in our simulations *a priori*, and must resort to simple assumptions anyway. An investigation of more detailed prescriptions for star formation efficiency in models for enrichment in the Milky Way is an interesting which we must reserve for future work due to how we have constructed the simulations presented in this paper.

- Motivated by these recent results suggesting nuance within the star formation law (Leroy et al. 2013; Ellison et al. 2020a; Kennicutt & de los Reyes 2020), we implement in our simulations a three-component power-law based on the results presented in Krumholz

et al. (2018). They construct a turbulence and mass-transfer driven theoretical model for star formation in disc galaxies, and present a comparison of its predictions under a handful of parameter choices to the data from Bigiel et al. (2010) and Leroy et al. (2013) (see their Fig. 2). We find that the following by-eye fit is a reasonable description of the aggregate data:

$$\dot{\Sigma}_\star \sim \Sigma_{\text{gas}}^N \quad (16a)$$

$$N = \begin{cases} 1.0 & (\Sigma_{\text{gas}} \geq 2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-2}) \\ 3.6 & (5 \times 10^6 \text{ M}_\odot \text{ kpc}^{-2} \leq \Sigma_{\text{gas}} \leq 2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-2}) \\ 1.7 & (\Sigma_{\text{gas}} \leq 5 \times 10^6 \text{ M}_\odot \text{ kpc}^{-2}) \end{cases} \quad (16b)$$

- The apparent linearity of the relationship above $\sim 2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-2}$ suggests that in this regime, star formation is proceeding at the fastest possible rate, and that $\tau_\star = \Sigma_{\text{gas}}/\dot{\Sigma}_\star = \text{constant}$. The results of Leroy et al. (2008, 2013) and Kennicutt & de los Reyes (2020) would suggest that this is the regime in which $f_{\text{mol}} = 1$. We therefore adopt the assumptions that f_{mol} is unity above an ISM surface density of $\Sigma_{\text{gas}} = 2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-2}$ and that τ_\star assumes that value of τ_{mol} at these densities, increasing with decreasing f_{mol} .

- With the formulation of the Kennicutt-Schmidt Law described by equation (16b) and the mathematical definition $\tau_\star \equiv \Sigma_{\text{gas}}/\dot{\Sigma}_\star$, we can derive $\tau_\star - \Sigma_{\text{gas}}$ or $\tau_\star - \dot{\Sigma}_\star$ relationships as required by VICE. With $\tau_\star = \tau_{\text{mol}}$ at $\Sigma_{\text{gas}} \geq 2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-2}$ setting the overall normalization, the following describes the resultant $\tau_\star - \Sigma_{\text{gas}}$ relationship:

$$\tau_\star = \begin{cases} \tau_{\text{mol}} & (\Sigma_{\text{gas}} \geq \Sigma_{\text{gas},1}) \\ \tau_{\text{mol}} \left(\frac{\Sigma_{\text{gas}}}{\Sigma_{\text{gas},1}} \right)^{\beta_1} & (\Sigma_{\text{gas},2} \leq \Sigma_{\text{gas}} \leq \Sigma_{\text{gas},1}) \\ \tau_{\text{mol}} \left(\frac{\Sigma_{\text{gas},2}}{\Sigma_{\text{gas},1}} \right)^{\beta_1} \left(\frac{\Sigma_{\text{gas}}}{\Sigma_{\text{gas},2}} \right)^{\beta_2} & (\Sigma_{\text{gas}} \leq \Sigma_{\text{gas},2}) \end{cases} \quad (17)$$

where we adopt $\Sigma_{\text{gas},1} = 2 \times 10^7 \text{ M}_\odot \text{ kpc}^{-2}$, $\Sigma_{\text{gas},2} = 5 \times 10^6 \text{ M}_\odot \text{ kpc}^{-2}$, $\beta_1 = -2.6$, and $\beta_2 = -0.7^2$. The additional factor of $(\Sigma_{\text{gas},2}/\Sigma_{\text{gas},1})^{\beta_1}$ simply ensures that this formalism is piecewise-continuous.

- Based on the observed Kennicutt-Schmidt relation at different redshifts, Tacconi et al. (2018) suggest that τ_{mol} should scale with redshift z and the deviation from the star forming main sequence δMS via $\tau_{\text{mol}} \propto (1+z)^{-0.6}\delta\text{MS}^{-0.44}$. We don't take into account the effect of δMS in our models, but we do investigate the redshift dependence. A reasonable approximation to the $t - z$ relation out to $z \approx 3$ assuming typical ΛCDM cosmology is given by:

$$\frac{t}{t_0} \approx (1+z)^{-5/4} \quad (18)$$

where t_0 is the present-day age of the universe, and t is not simulation time but the age of the universe at any given redshift. Plugging this relation into the Tacconi et al. (2018) scaling yields the following time-dependence for τ_{mol} :

$$\tau_{\text{mol}} = \tau_{\text{mol},0}(t/t_0)^{12/25} \approx \tau_{\text{mol},0}(t/t_0)^{1/2} \quad (19)$$

where $\tau_{\text{mol},0}$ is simply τ_{mol} at the present day. We generalize this formula to the following form:

$$\tau_{\text{mol}} = \tau_{\text{mol},0}(t/t_0)^\gamma \quad (20)$$

In this paper we present simulations which adopt $\tau_{\text{mol},0} = 2$

² $\beta = 1 - N$, where N is the power-law index of the $\dot{\Sigma}_\star - \Sigma_{\text{gas}}$ relationship.

Gyr (Leroy et al. 2008, 2013; Tacconi et al. 2018) and $\gamma = 1/2$ based on this argument. We have also ran simulations which adopt $\tau_{\text{mol},0} = 1$ Gyr and $\gamma = 0$ (a time-independent τ_{mol}), and found similar results.

- Fig. 6 shows τ_\star as a function of R_{gal} at six different time stamps predicted by our fiducial, inside-out SFH model. Wherever τ_\star is at its minimum value is where the ISM in that annulus is fully molecular.

– In the fiducial inside-out model, the molecular fraction is near unity out to $R_{\text{gal}} \approx 6$ kpc at all times.

– From the actual value of τ_\star at a given radius and time and the assumed τ_{mol} in the model, we can derive the model predicted molecular fractions within the disk via $f_{\text{mol}} = \tau_{\text{mol}}/\tau_\star$, then subsequently the HI + HII and H₂ masses from f_{mol} and the total gas supply after subtracting 27% to account for helium.

– These post-processing calculations suggest that our fiducial inside-out model predicts an HI + HII mass of $5.54 \times 10^9 M_\odot$ and an H₂ mass of $7.14 \times 10^9 M_\odot$ within $R_{\text{gal}} = 20$ kpc, the maximum radius in our model. This predicts a total H mass of $1.27 \times 10^{10} M_\odot$, and a total ISM mass of $1.74 \times 10^{10} M_\odot$ after adding 27% back to account for helium. This yields a global molecular fraction of $f_{\text{mol}} = H_2 / (HI + HII + H_2) = 56\%$. In comparison with observed results, this is an overprediction of f_{mol} ; Kalberla & Kerp (2009) report a value of 8×10^9 and $2.5 \times 10^9 M_\odot$ for the HI and H₂ masses within 60 kpc, the bulk of which is in the disk. They report an additional $\sim 2 \times 10^9 M_\odot$ in the warm ionized medium (HII). Heyer & Dame (2015) argue for a lower H₂ mass of $1 \times 10^9 M_\odot$. Our fiducial model is overpredicting the H₂ mass by a factor of ~ 3 and under-predicting the HI + HII mass by a factor of ~ 2 .

– Whatever the source of this discrepancy may be, we present in Appendix B a variation of our fiducial model with a significantly different assumption about τ_\star , and demonstrate that this decision does not impact our general conclusions. We reserve discussion on potential sources of this discrepancy to Appendix B as well. **This discussion may move back to this section, but I had some ideas on this, so I moved stuff later so that the methods section doesn't get lost in minutia.**

2.7 Summary

• In summary, our fiducial model has an inside-out SFH (see §2.4) with the star formation law as described in §2.6, radial migration the proceeds according to the diffusion model (see § 2.2), and yields and outflows as described in § 2.3.

• We have also conducted runs with the three other SFHs, the three other migration prescriptions, and the three other SFE prescriptions - a total of 64 simulations, as well as a variety of other test cases. In this paper, we present results wherever the model predictions are sensitive to the assumptions. However, in general the differences can be understood with only the variations in star formation history and the qualitative notion that many stars have moved beyond their birth radius.

• To ensure that resolution does not affect our results, we ran the same set of models with $n = 2$ stellar populations per zone per timestep, and found similar predictions.

3 METALLICITY SPACE

- Fig. 7 shows a scatter plot of 10,000 randomly sampled stellar populations in five bins of R_{gal} and three bins of $|z|$ ($R_{\text{gal}} = 3 - 5$

kpc, 5 - 7 kpc, 7 - 9 kpc, 9 - 11 kpc, and 11 - 13 kpc; $|z| = 0 - 0.5$ kpc, 0.5 - 1 kpc, and 1 - 2 kpc). These are the same bins and same scheme for organizing the panels as in Fig. 4 of Hayden et al. (2015).

• The width of the low- α sequence predicted by the model comes from radial migration. Though this is somewhat guaranteed by our model in choosing equilibrium abundances that reflect a realistic metallicity gradient, it is in good agreement with Schönrich & Binney (2009) and Nidever et al. (2014).

• The low- α sequence shifts from a high [Fe/H] locus at small R_{gal} to low [Fe/H] at high R_{gal} , in agreement with the observed distributions in APOGEE presented in Hayden et al. (2015).

• High- α stars are most prevalent at low R_{gal} and high $|z|$, and conversely for the low- α stars, also in agreement with Hayden et al. (2015).

– Similar results are found for different SFHs. This suggests that this observed result is a natural consequence of stellar migration.

– Only minor difference worth noting is that the starburst models predict a slightly higher characteristic [O/Fe] ($\sim +0.1$) for the low- α sequence. This is a natural consequence of the starburst producing young, α -enhanced stars (Johnson & Weinberg 2020).

• Left-hand panel of Fig. 8 compares the tracks predicted by our fiducial, inside-out SFH assuming diffusion migration (sudden) versus post-processing (dotted) for the gas-phase of a handful of radii denoted by the legend. Predicted [O/Fe]-[Fe/H] tracks for the diffusion model show significant deviations from the post-processing model. We demonstrate here that this is due to variability in the SN Ia rate induced by the time-dependent radial migration of the diffusion model. Simulation times of 2, 4, 6, 8, 10, and 12.7 Gyr shown in points and X's for the two models.

– For each zone, VICE provides in its outputs the rates of infall and star formation, the mass of the ISM, and the relevant abundance information for each element along with the associated MDFs at the final timestep. To determine the SN Ia rates, we therefore have to approximate from the output.

– The time-derivative of the mass of Fe in a given annulus is given by:

$$\dot{M}_{\text{Fe}} \approx y_{\text{Fe}}^{\text{CC}} \dot{M}_\star + y_{\text{Fe}}^{\text{Ia}} \langle \dot{M}_\star \rangle_{\text{Ia}} - \frac{M_{\text{Fe}}}{M_g} \dot{M}_\star (1 + \eta(R_{\text{gal}}) - r) \quad (21)$$

where this is an approximation because in detail, there is a small contribution from AGB stars, and the recycling in the simulation is done continuously, whereas here we simply take $r \approx 0.4$ (appropriate for a Kroupa IMF; Weinberg et al. 2017). This equation can be derived from the Weinberg et al. (2017) analytic models assuming CCSN and SN Ia enrichment for Fe with instantaneous recycling of previously produced Fe. VICE's science documentation could also be referenced here; it has a nice detailed section on its treatment of each term in handling enrichment rates³. Rearranging this for the term describing the rate of injection due to SNe Ia events, and normalizing by M_{Fe} yields the following proxy with units of frequency:

$$\frac{y_{\text{Fe}}^{\text{Ia}} \langle \dot{M}_\star \rangle_{\text{Ia}}}{M_{\text{Fe}}} \approx \frac{\dot{M}_{\text{Fe}}}{M_{\text{Fe}}} - y_{\text{Fe}}^{\text{CC}} \frac{\dot{M}_\star}{M_{\text{Fe}}} + \frac{\dot{M}_\star}{M_g} (1 + \eta(R_{\text{gal}}) - r) \quad (22)$$

This term on the left-hand side can be substituted with

³ https://vice-astro.readthedocs.io/en/latest/science_documentation/enrichment/index.html

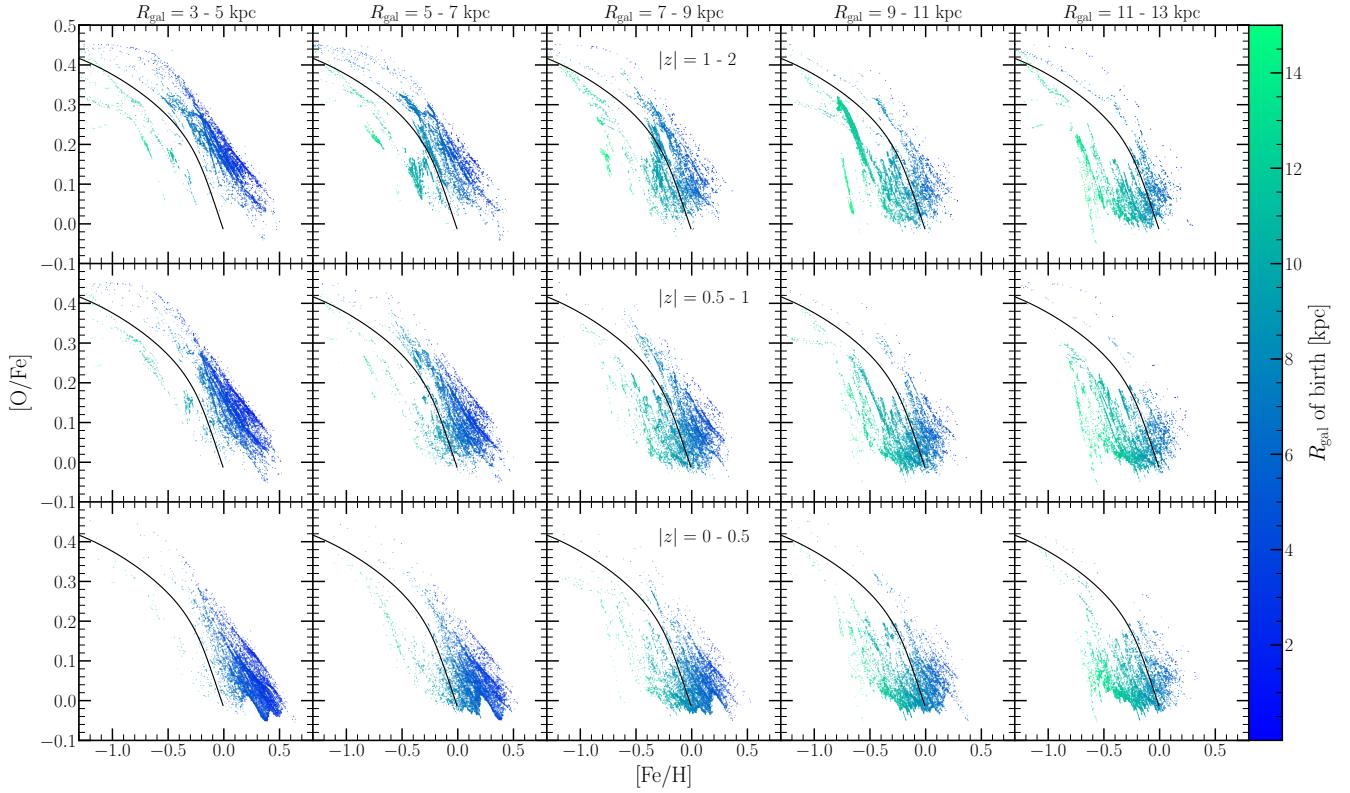


Figure 7. [O/Fe]-[Fe/H] diagrams for 15 galactic regions spanning five bins in R_{gal} and $|z|$. Each region has its own panel, with radial bins shown in columns denoted at the top of the figure, and with $|z|$ bins shown in rows denoted in text in the middle column. In each panel, we plot $N = 10,000$ points sampled from our simulated stellar populations in each region predicted by our inside-out SFH, where the probability of sampling is proportional to the present-day mass of each stellar population. In all panels points are color-coded according to the Galactocentric radius of birth of the stellar population. For reference, we plot in a solid black line in all panels the gas-phase [O/Fe]-[Fe/H] track predicted by the same SFH in the $R_{\text{gal}} = 8$ kpc annulus, but with the post-processing migration model; this curve is the same in all panels.

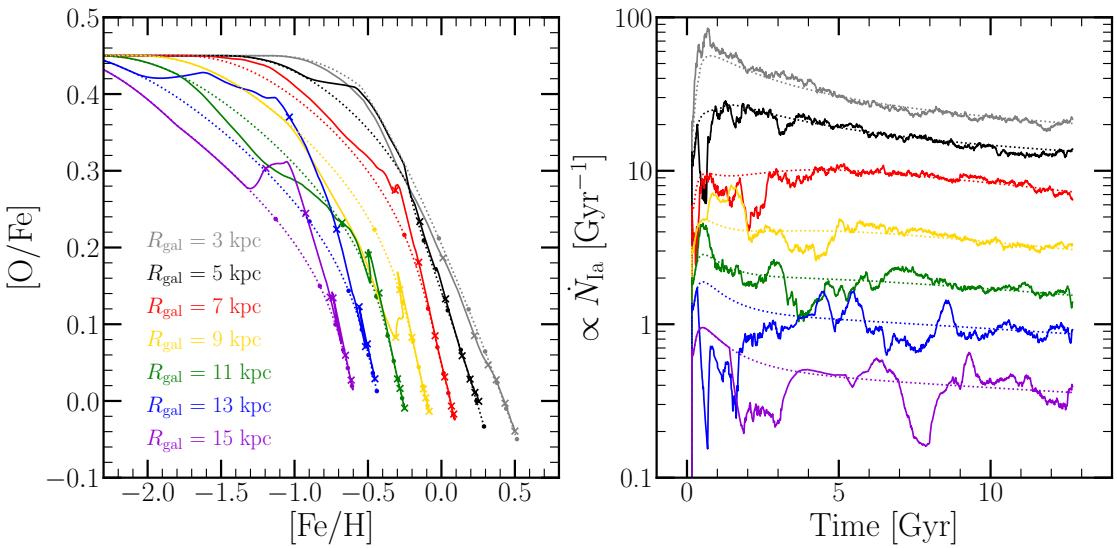


Figure 8. Left: Gas phase evolutionary tracks in the [O/Fe]-[Fe/H] plane for our inside-out SFH with either post-processing (dotted lines) or diffusion (solid lines) migration models. We plot tracks for seven annuli, color-coded according to their Galactocentric radius and denoted by the legend in the lower-left. We mark simulation times of 2, 4, 6, 8, 10, and 12.7 Gyr in X's for the diffusion model and points for the post-processing model. **Right:** The proxy for the SN Ia rate defined in equation (22) as a function of simulation time for the same annuli as in the left-hand panel. We multiply rates at each radii here by various prefactors to improve of clarity.

$m_{\text{Fe}}^{\text{Ia}} \dot{N}_{\text{Ia}} / M_{\text{Fe}}$, where $m_{\text{Fe}}^{\text{Ia}}$ is the average mass of Fe produced by a single SN Ia event, and \dot{N}_{Ia} is the SN Ia rate itself. For this reason, this equation constitutes a straight-forward proxy for the SN Ia rate at any given time.

– This proxy is plotted against simulation time in the right-hand panel of Fig. 8 for the same annuli as in the left-hand panel, with multiplicative factors added for visual clarity, diffusion model again shown in solid lines, post-processing in dotted lines. Whenever and wherever there is a deficit in SN Ia events relative to the post-processing model, the diffusion model tends toward higher [O/Fe] values than the post-processing scenario. Conversely, lower [O/Fe] for an excess in SN Ia events.

– SN Ia rates show high-amplitude variability on Gyr timescales, with low-amplitude white-noise on shorter timescales, potentially introduced at least in part by our discretization of the disc into annuli and the evolution into timesteps. The log-scaled y-axis makes it clear that the fractional amplitude is higher near the outskirts of the disc. This makes physical sense, because the stellar number density is much lower there, and as such would be much more susceptible to sampling noise - that is, a single star migrating has a larger fractional impact on the stellar density and thus the supernova rates at large radii than small radii.

- This is proof of concept that radial migration of nucleosynthetic yields can occur alongside stellar migration for delayed sources. In principle, it is reasonable to expect similar effects for s-process elements like carbon, nitrogen, strontium, yttrium, zirconium, etc. which are produced in AGB stars, though we do not investigate the impact for these elements here.

- In the literature, it's not uncommon to adopt tracks for different Galactocentric radii in the [O/Fe]-[Fe/H] plane to infer birth radii for observed stars. While an additional dimension such as age could mitigate these potential issues, we caution against inferring birth radii in this manner, because the tracks themselves appear to be sensitive to the dynamical history of the Galaxy. Based on the left-hand panel of Fig. 8, this effect could bias the inferred radii by many kpc. This however doesn't appear to be an issue for the low- α sequence, where the width of the [Fe/H] distribution appears to come from migration itself, as it did in Schönrich & Binney (2009) and Nidever et al. (2014). We therefore recommend that such a technique be restricted to $[\text{O}/\text{Fe}] \lesssim +0.1$ stars.

- Demonstrate in § 6.1 that this is a means with which to form α -rich and α -poor stars - or rather Fe-poor and Fe-rich, respectively.

4 METALLICITY GRADIENT

- Scaling of η with R_{gal} based on reproducing the observed mode($[\alpha/\text{H}]$)- R_{gal} trend, neglecting radial migration (see § 2.3). The target gradient is shown in a solid black line in the top panels of Fig. 9.

- In the top panels, red and blue stars show the mode O and Fe abundances in each annulus, and the shaded region shows the 16th and 84th percentiles of the MDF in that zone. Solid lines show the gas-phase abundance gradient at the present day. Bottom panels show the same thing for [O/Fe].

- Gradient is indeed recovered in [O/H], and radial migration appears to only induce scatter. While Fe did not enter into our procedure for setting the metallicity gradient, the model predicts a similar gradient for [Fe/H].

- Clear that the MDF shows a metal-rich mode and skew-

negative shape in the inner galaxy for both O and Fe. α -enhanced tail there as well. We demonstrate in § 4 that the MDF does shift to skew-positive in the outer galaxy, though this isn't as visually obvious from this plot due to the scatter in the mode at these radii.

- Details of the [O/Fe] gradient seem to be sensitive to differences in our SFHs, especially in the inner galaxy.

- Constant SFH is the only model in which the present-day gas phase gradient matches the stellar gradient at all radii. In remaining models, the present-day gas-phase abundance is above the majority of the stars because the star formation rate has decreased. We therefore expect observational studies of abundance gradients to differ noticeably between gas and stars.

- Stellar gradient is somewhat shallower in the late-burst model; this is because of the dilution associated with the starburst. Target gradient represents the equilibrium abundance at all radii, and we deliberately perturbed it from equilibrium, so any deviations from the expectation are a consequence of that.

- Late-burst model has super-equilibrium gas phase abundance at the present day. Can be seen by comparing it to the outer-burst model's gas phase gradient and seeing that it has a break at $R_{\text{gal}} = 6$ kpc, the threshold for the late starburst in this model. This is a consequence of the starburst as well - in infall driven starbursts, re-enrichment can produce super-equilibrium abundances which then decay back to the equilibrium abundance as the star formation rate declines (Johnson & Weinberg 2020).

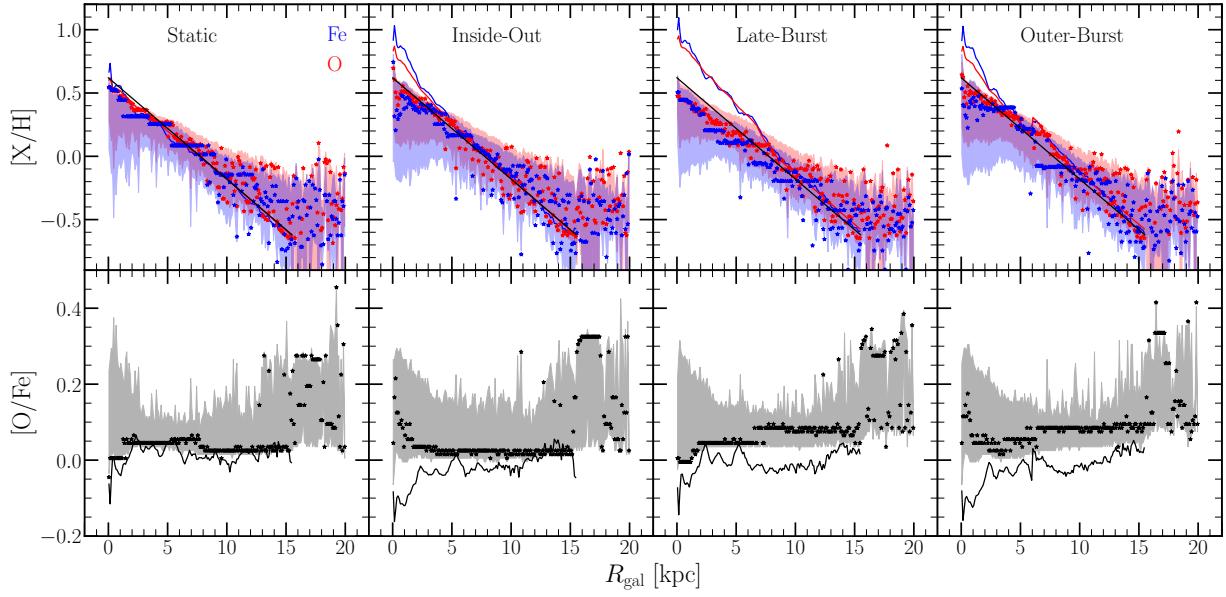


Figure 9. Radial abundance gradients in $[O/H]$ (top, red) $[Fe/H]$ (top, blue), and $[O/Fe]$ (bottom) for our four fiducial SFHs - constant (far left), inside-out (left-middle), late-burst (right-middle), and outer-burst (far right). We plot the gas phase abundance at the present day as a function of Galactocentric radius in solid lines. Stars denote the mode of the stellar MDF of the 100-pc width annulus at a given radius, with shaded regions marking the 16th and 84th percentiles thereof. Black lines in the top panels denote our target $[\alpha/H]$ gradient of mode($[\alpha/H]$) = +0.3 at $R_{\text{gal}} = 4$ kpc with a slope of -0.08 kpc^{-1} .

5 METALLICITY DISTRIBUTION FUNCTIONS

5.1 $[O/H]$ and $[Fe/H]$

- MDFs in bins of Galactocentric radius are a fundamental observable to test the validity of any chemical evolution model. In this section we compare our predicted MDFs to those observed in the 16th data release (DR16; Ahumada et al. 2020) of the Apache Point Observatory Galaxy Evolution Experiment (APOGEE; Majewski et al. 2017). The data is reduced using the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP; Holtzman et al. 2015; García Pérez et al. 2016). For further details on the APOGEE survey, a brief summary can be found in § 2 of Weinberg et al. (2019).

- We restrict our sample to stars with effective temperatures of $4000 \text{ K} \leq T_{\text{eff}} \leq 4600 \text{ K}$, surface gravities of $1.0 \leq \log g \leq 2.5$, and signal-to-noise ratios of at least 100. These cuts ensure that our sample consists of stars on the upper red giant branch, safely excluding red clump stars to avoid obvious systematics in the abundance distributions.

- Previously known that the MDFs in the disc midplane as observed in APOGEE show mode $[\alpha/H]$ and $[Fe/H]$ abundances that depend on Galactocentric radius, with a skew-negative distribution in the inner Galaxy and a skew-positive distribution in the outer Galaxy. Off the midplane, the MDFs merge and converge on $[\alpha/H] \approx [Fe/H] \approx -0.5$ (Hayden et al. 2015; Weinberg et al. 2019). This result is replicated for the observations in the right-hand column of panels in Fig. 10 and Fig. 11.

- Similar mode $[O/H]$ and $[Fe/H]$ between the 3 - 5 and 5 - 7 kpc in the APOGEE observations. What could be the origin of this? Cessation of star formation in the inner Galaxy? (see Fig. 1 of Peek 2009 and Fig. 2 of Fraternali & Tomasetti 2012). This would imply very few stars formed in the most metal-rich regions of the Galaxy, cutting off the MDF at high $[O/H]$, $[Fe/H]$. This

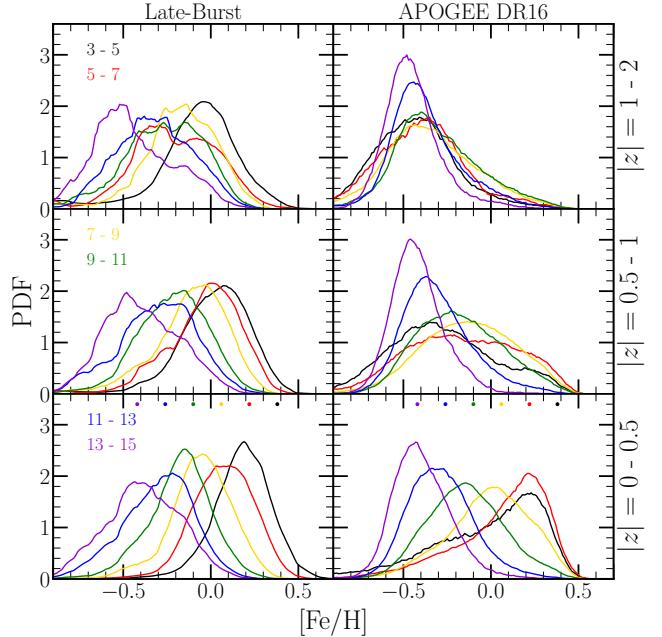


Figure 10. Metallicity Distribution Functions in $[Fe/H]$ predicted by our late-burst model (left) and as observed in APOGEE DR16 (right), for stars and simulated stellar populations with present day $|z| = 0 - 0.5$ kpc (bottom), $0.5 - 1$ kpc (middle), and $1 - 2$ kpc (top). MDFs are shown in bins of Galactocentric radius: 3 - 5 kpc (black), 5 - 7 kpc (red), 7 - 9 kpc (yellow), 9 - 11 kpc (green), 11 - 13 kpc (blue), and 13 - 15 kpc (purple). The points near the top of the bottom panels denote what the mode abundance would be if it followed our target gradient of $[Fe/H] = +0.3$ at $R_{\text{gal}} = 4$ kpc and a slope of -0.08 kpc^{-1} , assuming the inner radius of each bin (i.e. there is no point plotted for 15 kpc). All distributions are smoothed with a box-car width of $[Fe/H] \pm 0.1$.

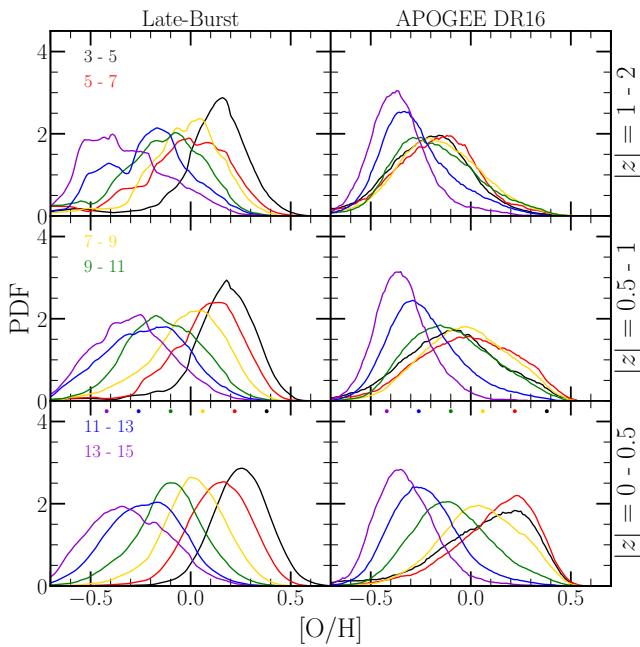


Figure 11. The same as Fig. 10, but for [O/H].

would suggest that the Milky Way has begun the quenching of star formation, a process believed to begin the centres of galaxies at this mass (e.g. Ellison et al. 2020b).

- Left-hand panels of Figs. 10 and 11 show the distributions predicted by our late-burst model. They successfully replicate the qualitative result that the mode [X/H] varies with present-day Galactocentric radius, but fail to show a similar mode abundance between the 3 - 5 and 5 - 7 kpc bins. Potentially linked to the cessation of star formation in inner Galaxy (Peek 2009; Fraternali & Tomassetti 2012), not included in our models.

- Beyond the midplane, our models fail to fully replicate the convergence of the MDFs at $[X/H] \approx -0.5$. The observed MDF at small radii shifts from skew-negative with a metal-rich mode to skew-positive with a metal-poor mode with increasing $|z|$. In our predicted MDFs for the inner galaxy, the mode does shift to lower $[X/H]$, though not to the same extent as in the observations. There is also very little change in skewness with $|z|$ predicted, in tension with observations.

– Could this point to a breakdown of our assumption that vertical mixing is efficient?

- We note that our models do a good job of producing a mode $[X/H]$ abundance in each radial bin close to our target gradient (shown as points plotted at the top of the bottom panels in Fig. 10 and 11 for the inner radius of each 2 kpc bin). In the inner galaxy bins, the predicted mode is moderately lower than the target, but this is due to the effect of dilution (see discussion in § 7). In our inside-out and outer-burst models, the difference in target and predicted mode $[X/H]$ is considerably smaller.

5.2 [O/Fe]

- In this section we compare our model predicted [O/Fe] distributions to those published in Vincenzo et al. (2021, in prep).

These are intended to simultaneously remove the effects of observational errors in [O/Fe] and the APOGEE selection function in these Galactic regions; that is, these are estimates of the *intrinsic* [O/Fe] distributions that, when convolved with observational uncertainties and the APOGEE selection function, would resemble the observed MDFs.

- Fig. 12 show distributions in [O/Fe] in two bins of [Fe/H] across 15 Galactic regions predicted by our inside-out SFH (solid lines). Dashed lines show the Vincenzo et al. (2021, in prep) distributions.

- At fixed R_{gal} and [Fe/H], the Vincenzo et al. (2021, in prep) MDFs show two peaks in the distribution which do not change with $|z|$; only their relative heights vary. This is an assumption built into the model, but the agreement with the APOGEE data is good. At small R_{gal} , the observed distributions may shift slightly to higher [Mg/Fe], but only at the $\lesssim 0.05$ level between $|z| \leq 0.5$ and $1 \leq |z| \leq 2.0$ (see their Fig. X). In contrast, our model predicted distributions show an increase in the mode [O/Fe] of ~ 0.1 over the same dynamic range in $|z|$. This suggests that our model overpredicts the increase of [O/Fe] with increasing $|z|$ compared to the APOGEE distributions. **As of now, I don't have a good explanation for why this is the case. Maybe the fact that we see this at small R_{gal} suggests the evolution of the bulge may have something to do with it? We're not modeling that here.**

- We note that the inside-out model fails to reproduce a Milky Way-like bimodality. Such a model prediction would appear as good agreement between the solid and dashed lines in Fig. 12, but that is not the case. There may be decent agreement in a given metallicity bin and Galactic region, but the agreement would need to be seen everywhere and at all metallicities.

- In the inner galaxy, we overestimate the [O/Fe] of the highest [Fe/H] stars at all $|z|$, and the differences in the distributions gets smaller with increasing R_{gal} . The lower [Fe/H] bins don't seem to have this problem.

- Since these are in specific bins in [Fe/H], this is an indication that our model is overpredicting the O abundances of these stars, rather than underpredicting [Fe/H].

- This could point to various things. Perhaps the Milky Way has different gradients in [O/H] than [Fe/H], an effect which is not captured by these models. Perhaps the innermost radii of the Milky Way has some contamination from bulge stars, another effect which is not captured by our models. Impact of bar evolution maybe?

- Similar results are found for our other SFHs.

- While the notion that an $[\alpha/\text{Fe}]$ dichotomy can arise out of radial migration alone was put forth in Schönrich & Binney (2009), and later explored by Nidever et al. (2014), this suggests that an inside-out star formation history combined with stellar migration is not conducive to predicting this observed result. The principle failure of this model is that it overpredicts the abundance of intermediate $[\alpha/\text{Fe}]$ stars. This could point to a handful of things.

- If the bimodality is to arise out of inside-out galaxy evolution and radial migration alone, the transition between low- and high- α sequences needs to occur faster than it does in these models. Applying the analytic model of Weinberg et al. (2017), this can be occur only with an adequately short τ_{\star} . Here we have adopted a star formation law which is motivated by observations (see discussion in § 2.6); this suggests that the observed rates of star formation are not conducive to an adequately short transition between the two sequences such as to predict the dichotomy in the models as it is observed in the Milky Way.

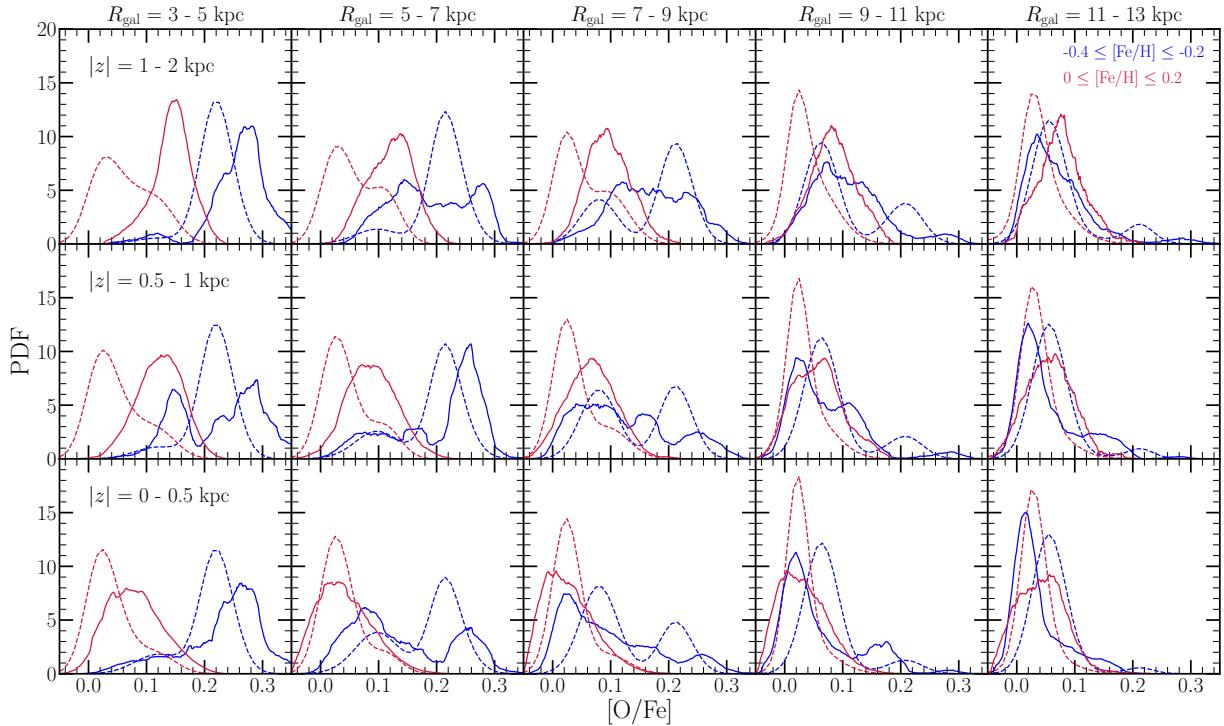


Figure 12. Predicted distributions in [O/Fe] in 15 Galactic regions and in two bins in [Fe/H]. Columns correspond to bins in R_{gal} , denoted at the top of each column. Rows correspond to bins in $|z|$, denoted in text in the left-hand column. Distributions are color-coded according to the [Fe/H] the sample is drawn from, denoted by the legend in the upper right panel. Solid lines represent that predicted by our inside-out SFH, while dashed lines correspond to the fits to the APOGEE DR16 data presented in Vincenzo et al. (2021, in prep), which quantify the intrinsic distributions accounting for observational uncertainties and the APOGEE selection function. Our simulated distributions are smoothed with a box-car width of $[\text{O}/\text{Fe}] \pm 0.02$ to improve clarity.

– Under our current assumptions, the simplest way to achieve this is to simply shut off star formation during the intermediate $[\alpha/\text{Fe}]$ phase in the simulations. This is indicative that two-infall evolutionary histories (e.g. Chiappini et al. 1997, 2001; Romano et al. 2010; Grisoni et al. 2017; Noguchi 2018; Spitoni et al. 2016, 2018, 2019, 2020) would improve the agreement between our model predictions and the observed distributions.

– Alternatively, there is observational evidence for a SN Ia Fe yield which scales inversely with metallicity (see discussion of Brown et al. 2019 in § 2.3). If this is the case, the decrease in [O/Fe] ratios at low [Fe/H] at the onset of the first SN Ia events may occur faster than in our models. This may predict the observed bimodality by ensuring that the transition between high- and low- α sequences is naturally fast due to supernova rates.

- This is at odds with the findings of Sharma et al. (2020), who claim to reproduce the $[\alpha/\text{Fe}]$ dichotomy with an analytic chemical evolution model with an inside-out SFH and stellar migration.

– We'll need to be careful in our final draft of the paper so that our language here isn't too strong. Below I've given my honest critiques of the Sharma et al. (2020) paper.

– They assume parameterized forms for [Fe/H] and $[\alpha/\text{Fe}]$ as functions of radius and time, chosen deliberately so that they agree with the observed data. Based on dynamical arguments similar to those in Schönrich & Binney (2009), their migration model has free parameters which then require fitting to observed data. Since the observed data is known to exhibit the dichotomy, their model's reproduction of the observed dichotomy is not a prediction but rather an assumption characterized a priori. For

this reason, rather than learning about the physical origins of the $[\alpha/\text{Fe}]$ dichotomy from Sharma et al. (2020), we instead get a robust mathematical characterization of the data.

– They find a dichotomy in the $[\alpha/\text{Fe}]\text{-}[Fe/\text{H}]$ plane to arise out of a fast transition in the gas phase between high- α and low- α . As we discussed here, we find in practice that this is only the case if τ_\star is sufficiently short, and our observationally motivated scaling would suggest that the required values are unphysically short ($\lesssim 1$ Gyr; we find that our $\tau_{\star,0} = 1$ Gyr model also fails to reproduce an adequately fast transition).

– We therefore argue that the claim of Sharma et al. (2020) that the observed dichotomy can arise purely out of inside-out galaxy formation with radial migration is likely contingent upon SN Ia rates scaling inversely with metallicity, as suggested by the findings of Brown et al. (2019). Under the parameters of our inside-out SFH model, this is the only physically realistic means with which the predicted transition between high- and low- α is adequately fast.

• These findings suggest that an inside-out SFH combined with radial migration, even when a late starburst is taken into account (as motivated by the findings of Isern 2019 and Mor et al. 2019), is *not* conducive to producing the observed $[\alpha/\text{Fe}]$ dichotomy. This suggests that more dramatic evolutionary events are likely responsible, such as a two-infall model (e.g. Chiappini et al. 1997, 2001; Romano et al. 2010; Grisoni et al. 2017; Noguchi 2018; Spitoni et al. 2016, 2018, 2019, 2020).

6 THE AGE-[α /FE] RELATION

- In this section, we compare our model predictions to the observational results of Feuillet et al. (2019). While we made use of APOGEE DR16 data in comparing our model predictions to the observed MDFs (Ahumada et al. 2020; Majewski et al. 2017), they make use of DR14 stars which have Gaia parallax measurements available (Abolfathi et al. 2018; Gaia Collaboration et al. 2018, for details on the APOGEE survey, see discussion in § 5). With their spatial and quality cuts, the final sample consisted of 77,562 stars.

- Feuillet et al. (2019) ages are measured via isochrone matching. Potentially give a little more detail, but that may be adequate for our purposes.

- In bins of [O/Fe], they assume a gaussian age distribution, and fit the mean and standard deviation to the observed sample. Because they assume a gaussian, they would report an equal mean and median. This is an important caveat in comparing our predicted relations to their results, because our model-predicted age distributions in bins of abundance are highly non-gaussian.

- The stellar populations from our simulations have different masses, so the age-distributions must be weighted by mass, since that scales with the number of stars that each stellar population represents. We therefore adopt a mass-weighted median age in bins of abundance as the appropriate comparison to the Feuillet et al. (2019) data. Physically, in a given bin [O/Fe], this is the age corresponding to the 50th percentile of the mass-weighted age distribution of our simulated stellar populations. For these reasons the comparison between our simulations and Feuillet et al. (2019) isn't exactly one-to-one. Potentially worth mentioning some of the systematics in calculating ages here as well, as that's relevant information.

6.1 The Impact of Radial Migration

- Fig. 13 shows a comparison between the predicted age-[α /Fe] relations in the solar annulus for our four migration models assuming our inside-out SFH.

- All models show reasonable agreement with the Feuillet et al. (2019) data; the population-averaged trend appears insensitive to the assumed migration model.

- Diffusion predicts the most intrinsic scatter, followed by linear, then sudden, then post-processing. This is a consequence of the variations in the SN Ia rates induced by time-dependent migration (see discussion in § 3). Further demonstration that under certain migration models, the radial migration of nucleosynthetic yields is statistically significant. This is also proof of concept that the effect is significant for abundance ratios of elements where at least one is produced by delayed nucleosynthetic sources. We therefore conclude that the time-dependence of radial migration is a necessary ingredient to chemical evolution models of galaxies where migration plays an important role, such as our own Milky Way. The level of scatter also appears to depend noticeably on which model for the time-dependence is adopted.

- This mechanism can produce populations of Fe-poor or Fe-rich stars, which can be misinterpreted as α -enhanced or α -deficient stars. Due to young stars migrating into or out of a given annulus, the SN Ia rate may be higher or lower than the expectation from a post-processing migration model. If this difference in the SN Ia rate is sustained for of order one depletion time, the ISM will be either Fe-poor or Fe-rich, and the stars that form there will inherit

that composition.⁴ The stars that form out of that patch of ISM can then migrate to the solar annulus. This effect is most significant at large Galactocentric radii where the fractional amplitude of the variability in the SN Ia rate is largest, and for that reason the young Fe-poor population predicted by our diffusion model originates at large radii ($\gtrsim 12$ kpc).

- Silva Aguirre et al. (2018) demonstrated that the observed young α -rich stars in the solar annulus have kinematics similar to the rest of the high- α population, and suggested that this may be the result of stellar mergers or mass transfer events, producing a population of truly old stars masquerading as young stars. This would imply that the observed young α -rich population is actually just older, high- α stars that have gone through some special class of stellar evolution. Our model predicts intrinsically young, Fe-poor stars to explain the observations, but these interpretations are not mutually exclusive. It's possible that some of the observed stars are truly young, Fe-poor stars, and that others underwent some stellar interaction event. Ascertaining the origins of this population therefore has implications for which of the migration models investigated here is the most realistic.

6.2 The Impact of the Star Formation SFH

- Fig. 13 shows a comparison between the predicted age-[α /Fe] relations in the solar annulus for our four SFHs assuming diffusion migration.

- Constant and inside-out SFHs describe the observed data the best. Both late starburst models show a population-averaged increase in [α /Fe] at young ages which is not observed in the data. This challenges the results of Isern (2019) and Mor et al. (2019), suggesting that these results on the Milky Way recent SFH are not consistent with chemical evolution models. If the Milky Way truly experienced a recent starburst, something not included in our models had to occur to prevent this global increase in [α /Fe].

- Below [O/Fe] $\approx +0.1$, the Feuillet et al. (2019) data seem to follow a slightly steeper age-[α /Fe] than our inside-out model predicts. This likely points to inaccuracies in the detailed form of the SFH or the SN Ia DTD, both of which are very plausible.

6.3 Beyond the Solar Annulus

- Fig. 14 presents a comparison of our simulation data to the Feuillet et al. (2019) observational data in 12 Galactic regions assuming the inside-out SFH.

- In the disc, the inside-out SFH is a reasonable description of the data for ages $\lesssim 5$ Gyr, above which the median ages are overpredicted relative to Feuillet et al. (2019). Far from the midplane, our model overpredicts the ages at nearly all abundances where Feuillet et al. (2019) have data, with the exception of the $R_{\text{gal}} = 7 - 9$ kpc and $|z| = 0.5 - 1$ kpc region.

- Differences in ages are interesting though - nearly everywhere we overpredict ages relative to Feuillet et al. (2019), their data are reasonably described by the stellar populations from our simulations that we would classify as Fe-poor (see discussion in § 6.1). Especially noticeable in the $R_{\text{gal}} = 5 - 7$ kpc, where the observed sample agrees nearly perfectly with one particular Fe-poor population in the $|z| \leq 0.5$ kpc and $0.5 \leq |z| \leq 1.0$ kpc regions. In the $|z| \leq$

⁴ Potentially note the Weinberg et al. (2017) definition: $\tau_{\text{dep}} \equiv \tau_{\star}/(1 + \eta - r)$. Even with τ_{\star} as high as ~ 5 Gyr at large Galactocentric radii, depletion times are still short due to the high values of η there.

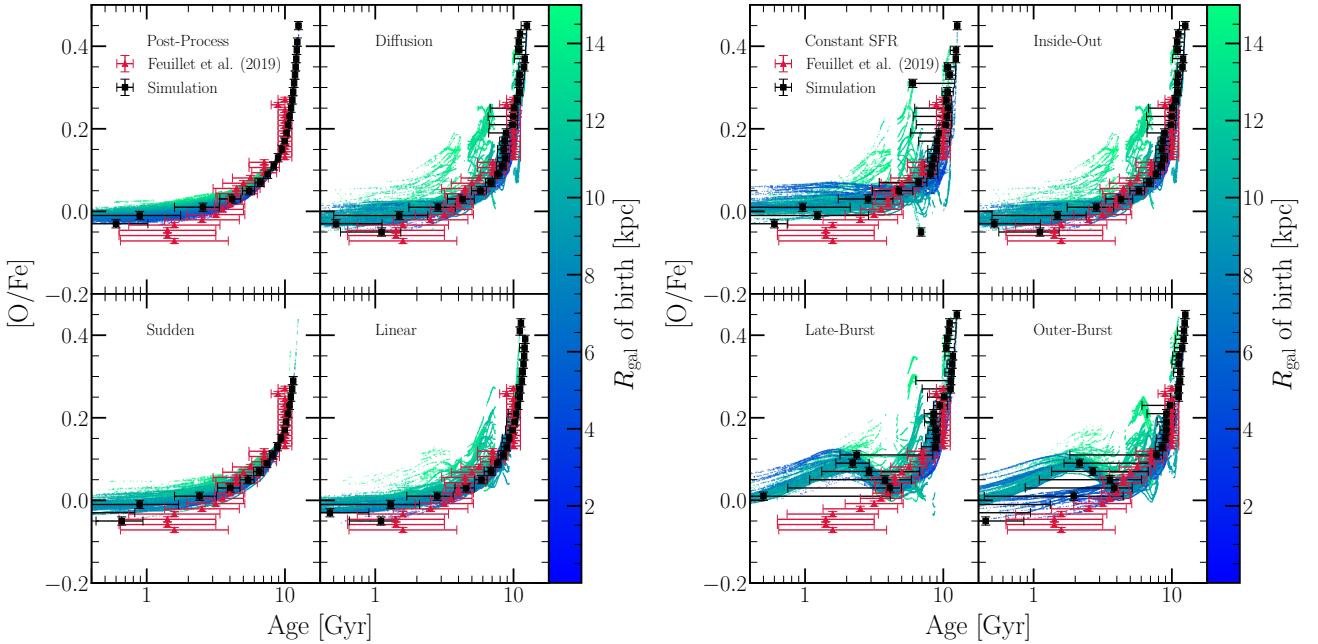


Figure 13. **Left:** A comparison of the predicted age-[O/Fe] relation for the solar annulus ($R_{\text{gal}} = 7 - 9$ kpc and $|z| \leq 0.5$ kpc) between the post-processing (upper left), diffusion (upper right), sudden (lower left), and linear (lower right) migration models, assuming our inside-out SFH. **Right:** The same as the left-hand panels, instead comparing the impact of our constant (upper left), inside-out (upper right), late-burst (lower left), and outer-burst (lower-right) SFHs. In all panels, red triangles and error bars denote the observed mean age and dispersion thereof in bins of [O/Fe] as reported by Feuillet et al. (2019); here we include only the bins containing at least 15 stars. Black squares denote the mass-weighted median age in 0.02-dex bins in [O/Fe] predicted by the simulations, with error bars denoting the 16th and 84th percentiles of the mass-weighted age distribution in those bins. Points in the background denote each individual stellar population from the simulation with a final position in the solar annulus, color-coded according to their Galactocentric radius of birth.

0.5 kpc population, the observed data even show an abrupt increase in age near the maximum [O/Fe] ratio of this particular subset of our stellar populations, and have one data point that agrees with the population-averaged trend in our simulation.

- Taking the Feuillet et al. (2019) data at face value, this would suggest that our simulation is overpredicting the rate of Fe injection from SN Ia, implying a SN Ia DTD whose characteristic timescales are longer than we employ here, thus slowing the decrease of $[\alpha/\text{Fe}]$ with time.

- Taking the simulation results at face value, this would suggest that APOGEE+Gaia target selection favors the stars we would classify as Fe-poor.

- This is another instance where a two-infall model would help improve the fit, however (e.g. Chiappini et al. 1997, 2001; Romano et al. 2010; Grisoni et al. 2017; Noguchi 2018; Spitoni et al. 2016, 2018, 2019, 2020). Through a perturbation of the ratio of CCSN to SN Ia rates (e.g. Johnson & Weinberg 2020), the model would predict an increase in $[\alpha/\text{Fe}]$ at ages of $\sim 8 - 10$ Gyr following the formation of the high- α sequence, decreasing toward younger ages. Although this would worsen the agreement in the solar annulus, it would improve it in other Galactic regions.

- We note that the intrinsic scatter in the age- $[\alpha/\text{Fe}]$ relation predicted by the model grows with increasing R_{gal} . Not only is the scatter in the colored background points visibly larger, but the error bars on the black points are as well, indicating that the age distribution is getting statistically significantly wider at fixed $[\alpha/\text{Fe}]$ with increasing R_{gal} .

- This is due to the variability in the SN Ia rates in each of our model’s annuli that we described in § 3. There, we demonstrated

that the highest amplitude variability in the SN Ia rate in our model galaxy is in the outskirts of the disc. This arises out of time-dependent migration inducing a variability; this variability gets larger in our models in regions where the stellar number density is low.

- With a higher variability in the SN Ia rate comes a higher variability in the rate of Fe production. Therefore, stars at large R_{gal} form out of an ISM with an [O/Fe] ratio that is more variable, the observational signature of which is the intrinsic scatter in the age- $[\alpha/\text{Fe}]$ relation growing with R_{gal} .

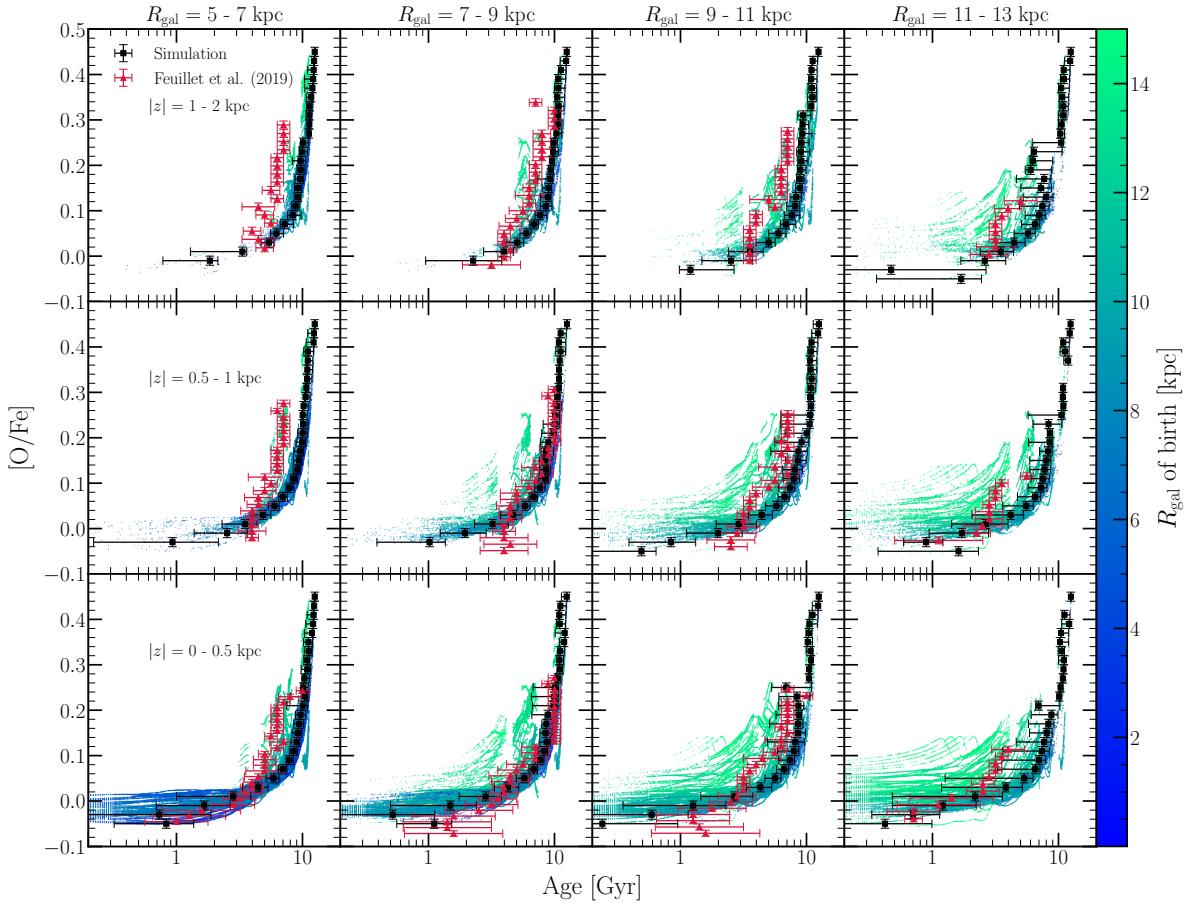


Figure 14. The age-[O/Fe] relation in 12 galactic regions predicted by our inside-out SFH. Bins in Galactocentric radius are shown in columns, and labeled at the top. Bins in the height $|z|$ above/below the disc midplane are shown in rows, noted in the left-hand column. Red triangles, black squares, error bars, and background points are as in Fig. 13 for the corresponding Galactic region.

7 THE AGE-METALLICITY RELATION

- Fig. 15 presents the age-[O/H] relation predicted by our constant-SFH model for the $|z| \leq 0.5$ kpc population at $R_{\text{gal}} = 5 - 7$, $7 - 9$, $9 - 11$, and $11 - 13$ kpc. The black points quantify the same mass-weighted median age in bins of [O/H] as in § 6. Colored points in the background are also the same - individual stellar populations color-coded according to birth radius.

- Feuillet et al. (2019) argued using the Weinberg et al. (2017) analytic models of chemical evolution that the non-monotonicity of the observed AMR was due to radial migration. The youngest stars at a given radius should reflect recent abundances of the local ISM, since young stars could not have migrated significant distances. Old stars, however, could have migrated from a wide range of radii, and can therefore reflect a wide range of abundances.

- Fig. 15 extends our understanding of this effect. In all regions plotted, the intrinsic scatter in the bulk age-[O/H] relation increases noticeably with increasing age, and the color-coding of the background points makes it clear that this arises out of radial migration.

- Furthermore, the characteristic metallicity of the youngest stars in a given radial bin decreases with increasing radius. This is a natural consequence of the abundance gradient that we've built in at late times, but interestingly, the effect is strong enough that at large radii, the median trend is nearly monotonically increasing with age. This is completely backwards from what is expected from one-

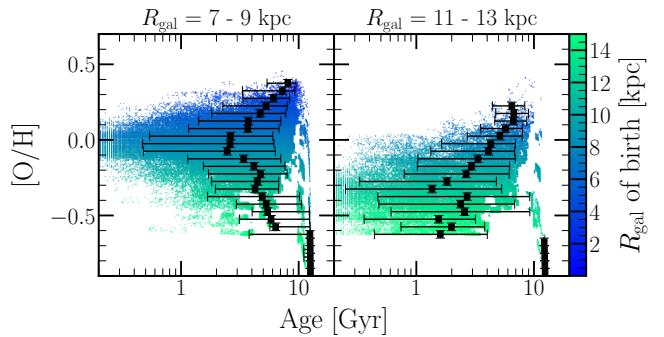


Figure 15. The age-[O/H] relation predicted by our constant SFR model for $R_{\text{gal}} = 7 - 9$ kpc (left) and $11 - 13$ kpc (right). Each panel plots only the $|z| \leq 0.5$ kpc population. The colored points in the background and the black squares with error bars are as in Fig. 13, but with our binned, simulation prediction quantified in 0.05-dex bins in [O/H].

zone models of chemical evolution for any radius. Since this is our constant-SFR model, it is not subject to the effects of a time-varying SFH, quantifying only what is caused by radial migration.

- Similar things are found for Fe, but the variations in the black squares a bit bigger so we demonstrate the effect with [O/H].

- Fig. 16 shows the age-[O/H] (top panel) and the age-[Fe/H]

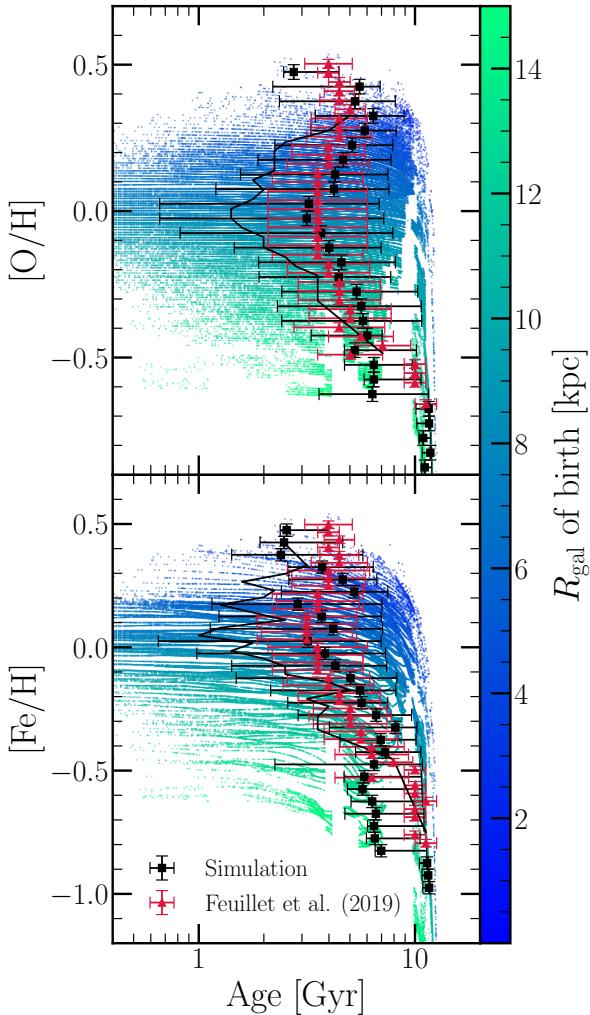


Figure 16. The age-[O/H] (top) and age-[Fe/H] (bottom) relations for the solar annulus (i.e. $R_{\text{gal}} = 7 - 9$ kpc, $|z| \leq 0.5$ kpc) as predicted by our inside-out SFH. Red triangles, black squares, error bars, and background points are as in Fig. 13, but with our simulation prediction quantified in 0.05-dex bins in [O/H] and [Fe/H]. For comparison, we plot the Feuillet et al. (2018) data in a solid black line, omitting the associated uncertainties for the sake of clarity.

(bottom panel) relations predicted by our fiducial inside-out SFH in the solar annulus ($R_{\text{gal}} = 7 - 9$ kpc and $|z| \leq 0.5$ kpc), under the same plotting convention as in § 6. We add a solid black line to denote the Feuillet et al. (2018) AMR in comparison to that of Feuillet et al. (2019).

– We note that we find similar results with regard to the age-[O/H] relation predicted by our models.

– Feuillet et al. (2018) shows a significantly younger median age at solar metallicity in both O and Fe ($\sim 1 - 2$ Gyr as opposed to $\sim 2 - 3$ Gyr).

– Fiducial, inside-out SFH performs decently at explaining the Feuillet et al. (2019) AMR, but the ages at solar metallicity in the Feuillet et al. (2018) study are too young for this model. Comparing the black points in this figure to those in Fig. 15 suggests that the constant SFR model can't explain ages that young either. If we were to take the Feuillet et al. (2018) results at face value in comparison to these models, this would strongly

suggest a recent enhancement at least in the local star formation rate of the solar annulus to increase the frequency of ~ 1 Gyr old, solar metallicity stars close to the sun.

– Such a burst is indeed supported by the findings of Mor et al. (2019), who find a factor of ~ 2 enhancement in the SFH of the Milky Way ~ 2 Gyr ago by comparing population synthesis models to observed stellar luminosity functions and color-magnitude diagrams with Gaia data (Gaia Collaboration et al. 2018). Isern (2019) reach similar conclusions modeling white dwarf luminosity functions in the solar neighborhood with Gaia parallaxes.

- Fig. 17 shows a comparison of the predicted age-[Fe/H] relation for $|z| \leq 0.5$ kpc stars in the $R_{\text{gal}} = 5 - 7$, $7 - 9$, $9 - 11$, and $11 - 13$ kpc annuli. Points are plotted in the same manner as in Fig. 15 and Fig. 16 in this section.

– Beyond the solar annulus, we note that our inside-out SFH model in general overpredicts the characteristic ages of stars compared to Feuillet et al. (2019) except for $[\text{Fe}/\text{H}] \approx +0.4$ and -0.5 . We also remark that in the trend is not very well reproduced either, particularly in the $R_{\text{gal}} = 5 - 7$ kpc bin.

– In the bottom row of panels, we show the same relation for our late-burst SFH. Interestingly, the late-burst model improves upon the failures of the inside-out model significantly. In particular, the over-prediction of ages of solar and intermediate metallicity stars is fixed by the late-burst model.

- In our recent starburst models, VICE calculates that there is a significant amount of zero metallicity gas infall required to sustain such an increase in star formation under our adopted $\dot{\Sigma}_{\star} - \Sigma_{\text{gas}}$ relation (see bottom two panels of Fig. 4 and § 2.6).

– Denoting the ages and compositions of the individual stellar populations from our simulations, the colored points in the background of these panels trace the metallicity of the gas as a function of time at various radii. That is, the blue points also represent the [Fe/H] of the gas phase at small radii at various lookback times, and the same for the green points and large radii.

- This demonstrates that at a lookback time of ~ 2 Gyr, by construction, the ISM at nearly all Galactocentric radii decreased in metallicity. At the same time, the star formation rates increased by a factor of ~ 2 , again by construction (see the right-hand panels of Fig. 4). The result is that at any radius, the frequency of $-0.5 \lesssim [\text{O}/\text{H}] \lesssim 0$ stars at ages of ~ 2 Gyr is increases by a factor of ~ 2 from the inside-out model, decreasing the characteristic ages of stars at these metallicities.

- Although there appears to be an offset between the Feuillet et al. (2019) data and our predicted AMR at high R_{gal} , the late-burst model describes the observed trend noticeably better than the inside-out model. Although the ages of modestly sub-solar metallicity stars in the solar annulus are slightly under-predicted by the model, the same can be said about the trend here and at $5 - 7$ kpc as well.

- The parameters which control this offset are the yields of Fe and the mass-loading factor as a function of R_{gal} . If we were to take a slightly higher yield of Fe, it would however increase the overall normalization at $R_{\text{gal}} < 9$ kpc as well, where we do not see this discrepancy. It's possible that the mass-loading factor η (see discussion in § 2.3) increases too quickly at large R_{gal} in our models, or that outflows are more efficient at

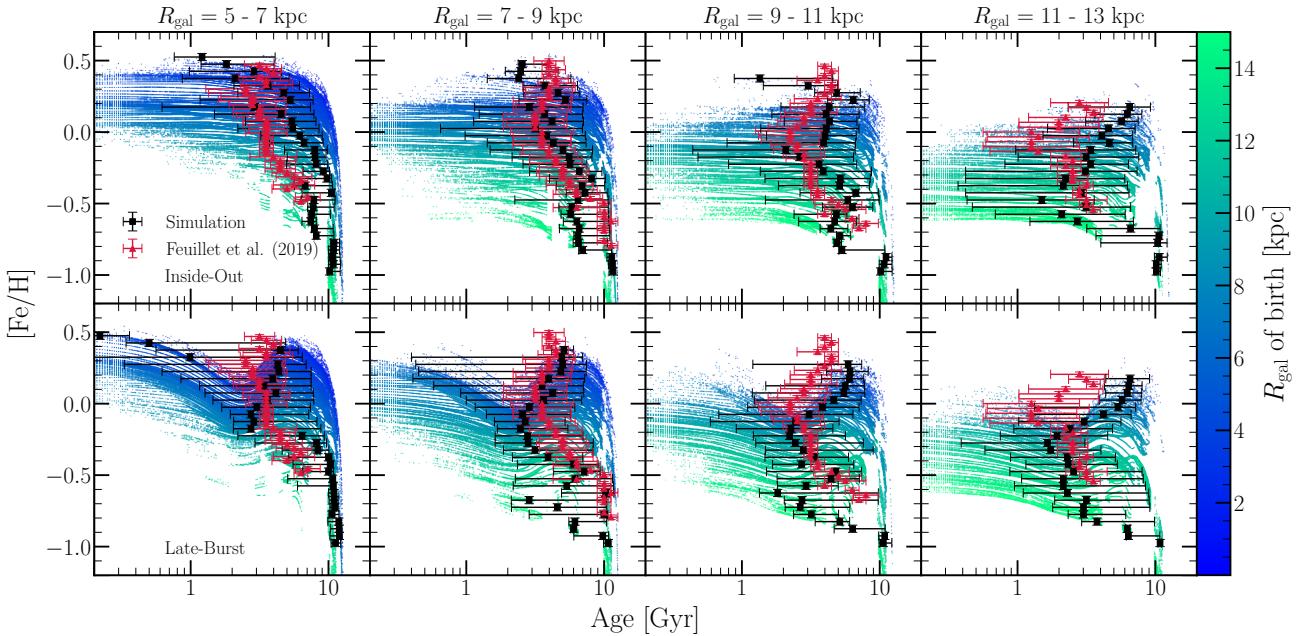


Figure 17. The age-[Fe/H] relation predicted by our inside-out (top) and late-burst (bottom) SFHs for $R_{\text{gal}} = 5 - 7 \text{ kpc}$ (left), $7 - 9 \text{ kpc}$ (left middle), $9 - 11 \text{ kpc}$ (right middle), and $11 - 13 \text{ kpc}$ (right). Each panel plots only the $|z| \leq 0.5 \text{ kpc}$ population. Red triangles, black squares, error bars, and background points are in Fig. 13 for the corresponding Galactic region, but with our simulation prediction quantified in 0.05-dex bins in [Fe/H].

removing individual elements from the star forming reservoir. Variable metal-loading factors η_Z would be supported by the observations of Chisholm, Tremonti & Leitherer (2018).

- We remark that the late-burst model predicts extremely young characteristic ages for the highest [Fe/H] bins in the 5 - 7 kpc annulus. This is merely a consequence of the re-enrichment that accompanies the ensuing starburst; in infall-driven starbursts such as this, the abundances often reach super-equilibrium values while the star formation rate is still perturbed, which then decay together back to their pre-burst, equilibrium values (Johnson & Weinberg 2020). On Fig. 17, this is noticeable by the abundances on the left-hand edge of each panel being higher in the late-burst scenario. The points that we have plotted as characteristic ages represent that 50th percentile of the mass-weighted age distribution of our stellar populations in a bin of abundance; in other words, a mass-weighted *median*. The age-distribution at these abundances is bimodal, noticeable by the colored points in the background, indicating that the median may not be the best statistic for these few stellar populations. We note that the outer-burst model, in which $R_{\text{gal}} < 6 \text{ kpc}$ do not experience the starburst, does not suffer from this issue. The minutia of the AMR at different radii in our late starburst models is sensitive to the detailed time-dependence of the starburst in each annulus.

- Conclude that the AMR reported by Feuillet et al. (2019) is better described by our late starburst models than the inside-out model. This is at odds with our findings in the same simulations with the same observational dataset with regard to the age-[α /Fe] relation. Perhaps the differences can be reconciled.

- The models would be able to have their cake and eat it too if there is a recent starburst with no ensuing α -enhancement.

This is difficult to rationalize, however, because an increase in [α /Fe] in the wake of a starburst is a direct consequence of the perturbed ratio of CCSN-SN Ia rates caused by the burst (Johnson & Weinberg 2020). We argue based on this that it is unlikely that the detailed timing of a recent starburst would mitigate this issue.

- It's possible that the Milky Way experienced dilution with no ensuing starburst. This could be the case if the accreted gas was mostly in the form of HI or HII that has not yet cooled and been available for star formation, but has been mixing with the nucleosynthetic products of ongoing star formation in the Galaxy. With dilution playing a noticeable role in the AMR predicted by our burst models, it's possible a model of this nature could agree with both the AMR and the age-[α /Fe] relation. This would require future studies which include a treatment of a multi-phase ISM.

- As a general result here, we caution future studies against leveraging the agreement between a chemical evolution model and observational data based on the solar annulus alone. Fig. 16 presents the comparison of our fiducial, inside-out SFH model predictions in the solar annulus to the Feuillet et al. (2019) data, and Fig. 17 compares the inside-out and late-burst predictions in a larger range of radii. Had we only considered the solar annulus, we would have concluded that the fiducial, inside-out model agrees with the Feuillet et al. (2019) data well. However, in considering other regions of the Galaxy, we found that the inside-out model actually had a handful of failures which were mitigated by our late-burst model.

- With our age-[O/H] and age-[Fe/H] relations, we find similar results as in § 6.3 whereby our model in general overpredicts the observed ages of stars at high $|z|$. Discussion of potential sources of this can be found in § 6.3.

8 CONCLUSIONS

- We have modeled the Milky Way as a series of concentric annuli with $\Delta R_{\text{gal}} = 100$ pc width, describing each annulus as a conventional one-zone model of chemical evolution, and allowing the exchange of stellar populations between zones to model the impact of stellar migration on enrichment in the Galaxy. The resulting “multi-zone” is a technique that has received attention from only a handful of chemical evolution studies to date (Matteucci & Francois 1989; Schönrich & Binney 2009; Minchev et al. 2013; Sharma et al. 2020). In so doing, we adopted star particles from the h277 zoom-in, hydrodynamical simulation to describe the dynamical history of a Milky Way-like galaxy at face value. This yields a model for radial migration which does not have any free parameters (see discussion in § 2.1).

- We found that similar numbers of star particles in h277 migrated inward as outward, and that the two proceed on different timescales, indicating that they’re potentially linked to different physical processes. We found that distributions of birth radii in bins of final radii showing a mode which depends strongly on age is a consequence of the surface density gradient. It’s not that stars in general migrate outward, it’s simply that there are many many stars born at small radii presently at large radii due to the fact that there are so many more stars born at small radii, a consequence of the steep nature of the surface density gradient (e.g. Bland-Hawthorn & Gerhard 2016).

- We found that the e-folding timescales for star formation indicated by population averaged observations of Milky Way like galaxies are long (~ 15 Gyr in the solar annulus; Sánchez 2020).

- We find that through stellar migration our models naturally reproduce the observed result that high- α stars are more common at small R_{gal} and high- $|z|$, and the opposite for low- α stars (Hayden et al. 2015; see Fig. 7 and associated discussion in §3).

- We demonstrate that the time-dependence of radial migration has a significant impact on the [O/Fe] ratios as a function of time and radius by inducing variability in the SN Ia rate and therefore the Fe abundance (see Fig. 8 and the associated discussion in §3). Although we haven’t investigated it here, we expect similar arguments to extend to other delayed sources such as AGB stars and the associated s-process elements such as carbon, nitrogen, strontium, yttrium, and zirconium. This is proof of concept that under certain models for the time-dependence of radial migration, the migration of nucleosynthetic yields proceeding alongside stellar migration is a statistically significant effect. We therefore recommend future studies relax the assumption that stars do not contribute to nucleosynthesis beyond their birth radius.

- We find that our models naturally reproduce the variations in the observed metallicity distribution functions (MDFs) with Galactocentric radius in the disc midplane ($|z| \leq 0.5$ kpc). At low R_{gal} , the mode abundance in both [O/H] and [Fe/H] is high, and conversely low at high R_{gal} , with skewness in qualitative agreement with the observations (e.g. Hayden et al. 2015; see Fig. 10 and Fig. 11 and the associated discussion in § 5.1). In some sense, however, this is more a consequence of our building a realistic abundance gradient into our models than a prediction on its own. In the observations, the mode abundance in the midplane at $R_{\text{gal}} = 3 - 5$ kpc is strikingly similar to that at 5 - 7 kpc. Our models fail to reproduce this, though it may be tied to a cessation of star formation in the inner galaxy, as suggested by Peek (2009) and Fraternali & Tomasetti (2012). Off the midplane, our models predict a slight shift of the MDFs at small R_{gal} to lower metallicity, where in the observations, those for all radii converge at $[\text{Fe}/\text{H}] \approx [\text{O}/\text{H}] \approx -0.5$.

- We find that none of our models predict a Milky Way-like dichotomy in $[\alpha/\text{Fe}]$, as observed by many studies. The principle failure is that the frequency of intermediate $[\alpha/\text{Fe}]$ stars is overpredicted relative to the observations (Hayden et al. 2015; Vincenzo et al. 2021, in prep). This suggests that inside-out galaxy growth combined with radial migration, even if a recent starburst is included (e.g. Mor et al. 2019; Isern 2019), is not adequate to explain the observed dichotomy. This is at odds with the findings of Sharma et al. (2020), though we believe our study to have a more physically motivated treatment of the chemical evolution models (see discussion in § 5.2). The over-prediction of intermediate $[\alpha/\text{Fe}]$ stars in general suggests that shutting off star formation for the intermediate- α phase would mitigate the problems suffered by our models; we therefore postulate that a two-infall model may be necessary to explain the observed dichotomy (e.g. Chiappini et al. 1997, 2001; Romano et al. 2010; Grisoni et al. 2017; Noguchi 2018; Spitoni et al. 2016, 2018, 2019, 2020). Alternatively, we speculate that the issues may be mitigated by an effective SN Ia yield of Fe which depends inversely on metallicity, based on the findings of Brown et al. (2019).

- We found that our recipe for setting the gradients in both stellar surface density and metallicity yields the correct result. This is an indication that radial migration does not change the profile of these observed trends, only inducing scatter.

- We found that slightly different predictions in the stellar and gas-phase metallicity gradient are natural consequences of variations in the SFH.

- We found that different models for the SFH predicted different [O/Fe] gradients at the present day.

- See Fig. 9 and the associated discussion in § 4 for the metallicity gradient; Fig. 5 and discussion in § 2.5 for the surface density gradient.

- We found that the observationally motivated approach to modeling the efficiency of star formation ($\tau_{\star} \equiv \Sigma_{\text{gas}}/\dot{\Sigma}_{\star}$) adopted in our simulations in general overpredicts the molecular fractions by mass of the model Galaxy. The discrepancy arose out of an underabundance of HI/HII rather than an overabundance of H₂ in the global demographics predicted by our simulation. Nonetheless, we demonstrate in Appendix B that relaxing these assumptions do not impact our model predictions. We also discuss a handful of potential sources of the discrepancy there.

- We found that the time-dependence of radial migration indeed does have an impact on enrichment. Previous studies in the literature have assumed that this is not the case, due to the slow nature of migration (e.g. Minchev et al. 2013). However, we find that even for young stars, the tails of the present-day R_{gal} distribution are adequately long that when age-dependent migration is taken into account, the model predicts variations in the SN Ia rate that then impact predicted abundance ratios. We demonstrated that this is a means with which stellar migration may produce young, Fe-poor stars which can migrate to the solar annulus, and potentially be misinterpreted as young, α -rich stars. This is proof of concept that the radial migration of nucleosynthetic yields can proceed alongside the radial migration of stars, depending on which model for migration that you believe.

- These stars have been seen in observed data from APOGEE (e.g. Silva Aguirre et al. 2018). They postulate that since these stars have kinematics similar to the rest of the high- α population, that perhaps they are the consequence of stellar mergers or mass transfer events, producing truly old stars that simply appear to be younger. While these interpretations are not mutually exclusive,

an observational test to ascertain the origins of the observed young α -rich population would have implications for which of the models for radial migration investigated here are the most realistic.

- In the observations, the high- α sequence is known to be most prominent at low R_{gal} and high $|z|$, and conversely, the low- α population at high R_{gal} and low $|z|$ (e.g. [Hayden et al. 2015](#)). We found that this is a natural consequence of stellar migration, though the detailed distributions depend on the SFH and the relative yields of α and iron-peak elements, both of which are highly uncertain, and we do not model them in detail here.

- We found that our inside-out model predicts an age-[α/Fe] relation which is in good agreement with the observed relation reported by [Feuillet et al. \(2019\)](#). Our recent starburst models predict a global α -enhancement that simply isn't seen in the data. This suggests that our current understanding of chemical evolution may be at odds with the findings of [Mor et al. \(2019\)](#) and [Isern \(2019\)](#) suggesting a recent factor of ~ 2 enhancement in the Milky Way SFH peaking ~ 2 Gyr ago.

- We find that where our recent starburst models fail to agree with the age-[O/Fe] relation reported by [Feuillet et al. \(2019\)](#), it tends to overpredict ages at a given [O/Fe]. However, it tends to fail in a manner such that the [Feuillet et al. \(2019\)](#) relation agrees with the stars that we would call Fe-poor - that is, the stars that are α -enhanced due to forming out of an Fe-poor ISM due to the same variations in the SN Ia rate that can produce young, α -rich stars. While this may be coincidence, if we take the observations at face value, it would suggest that our model is overpredicting the rate of Fe injection, and that perhaps SN Ia yields of Fe should be lower. If we take the model at face value, it suggests that APOGEE+Gaia targets preferentially sample young, α -rich stars ([Feuillet et al. 2019](#) used APOGEE data for stars with Gaia parallaxes).

- We found that the age-metallicity relation (AMR) in both [O/H] and [Fe/H] reported by [Feuillet et al. \(2019\)](#) is better fit by our late starburst models. This is at odds with the findings regarding the age-[O/Fe] relation, where the inside-out model agreed with the data much better than the starburst models. This is interesting, because different observables are favoring different models for the SFH, suggesting that something more complicated is going on.

– The starburst models agreeing with the observed AMR has to do with dilution as well as the starburst itself. It's possible that the Milky Way experienced a dilution event that was primarily HI/HII gas that has not yet cooled to be available for star formation. This could mitigate the discrepancy. It would, however, be at odds with the observed results of [Mor et al. \(2019\)](#) and [Isern \(2019\)](#) suggesting such a recent enhancement in star formation.

– Off the midplane, our model-predicted AMR overestimates ages relative to [Feuillet et al. \(2019\)](#), just like in the age-[O/Fe] relation. We speculate that this may be related to the Sagitarrius dwarf galaxy; it's important to remember that our model galaxy does not have the Milky Way's dynamical history, but h277's dynamical history. [Jon and I are looking into this, but as I understand it, what I've written next is in general true.](#) h277 did not have a Sagitarrius-like accretion event. Since Sagitarrius has made multiple pericentric passages nearly head-on with the Milky Way disc, it's possible that young stars were kinematically heated to high $|z|$, an effect that would not be present in h277 without such an event. This would decrease our model predicted median ages at these heights by directly adding more young stars, and potentially not having noticeable effect in the midplane due to the much larger number of stars there.

- We remark on the low number of multi-zone chemical evolution models in the literature. We call for more studies which adopt a similar approach; with only a handful of simulations which can be ran in a combined time interval of less than a single working day, we were able to assess model predictions of various chemical evolution scenarios in comparison to a wide range of observables. With a wealth of one-zone chemical evolution models (both numerical and analytic) and high-resolution hydrodynamical simulations already in the literature, there is a true void in the literature for these medium resolution, medium computational expense models which which can teach us a great deal about the enrichment history of the Milky Way. For this reason, VICE is open source software, and its `milkyway` object which ran our simulations adopts many of this paper's physically and observationally motivated assumptions as default values. Alternative zone configurations can be achieved by subclassing the `multizone` object and specifying how gas and stars should move between the individual zones, as we have already done for the `milkyway` object.

- VICE is publicly available and open-source. It can be installed via pip (<https://pypi.org/project/vice>). Documentation is available at <https://vice-astro.readthedocs.io>. Source code is hosted at <https://github.com/giganano/VICE.git>. Python code which runs the simulations presented in this paper are included as supplementary material in the `git` repository; our models can be ran directly from a bash terminal without modifying the source code, and are capable of predicting abundances for ~ 2 million stellar populations in only ~ 2 CPU hours with a single core on personal computers.

9 ACKNOWLEDGEMENTS

We are grateful to Diane Feuillet for sharing the data from [Feuillet et al. \(2018\)](#) and [Feuillet et al. \(2019\)](#) with us. There will be others added, depending on whether or not they go here or on the author's list. I'll also need to add the SDSS acknowledgements since we made use of APOGEE data.

Software: Matplotlib ([Hunter 2007](#)); Astropy ([Astropy Collaboration et al. 2013, 2018](#)); NumPy ([Harris et al. 2020](#)).

10 DATA AVAILABILITY

In case anyone hasn't seen one of these Data Availability statements yet, this is now a requirement by MNRAS. It wasn't when I submitted my last paper, but was by the time we were finished with the referee report, so I wound up having to add one. They just want you to say if the data are available to the reader or not, and where/how they can get it if they are. VICE is open source software, and as such the source code for these simulations is publicly available.⁵ The source code which produces the outputs presented in this paper as well as the figures are included as secondary material in the GitHub repository. While the aggregate of all outputs analyzed in this paper are sufficiently large that it is not conducive to store them on GitHub, we provide instructions on how to run our simulations and variations thereof. All observational data appearing in this paper is publicly available, and is also included with the source code for our simulations and figures.

⁵ <https://pypi.org/project/vice>
<https://vice-astro.readthedocs.io>
<https://github.com/giganano/VICE.git>

Appendices

A NORMALIZING A FIDUCIAL STAR FORMATION HISTORY

- Derive formula for normalizing an SFH given the time-dependence at a given radius $f(t|R_{\text{gal}})$ and the radial dependence of the desired surface density gradient at late times $g(R_{\text{gal}})$. Neither need be normalized.

$$\dot{\Sigma}_{\star}(R_{\text{gal}}, t) = \dot{\Sigma}_{\star,0}(R_{\text{gal}})f(t|R_{\text{gal}}) \quad (23)$$

$$\Sigma_{\star}(r) = \Sigma_{\star,0}g(R_{\text{gal}}) \quad (24)$$

- Integrate surface density of star formation with time and you get the present day surface density gradient at that radius. This yields the unknown $\dot{\Sigma}_{\star,0}$ in terms of $\Sigma_{\star}(R_{\text{gal}})$ and subsequently the unknown $\Sigma_{\star,0}$.

$$\Sigma_{\star}(R_{\text{gal}}) \approx (1 - r) \int_0^T \dot{\Sigma}_{\star}(R_{\text{gal}}, t) dt \quad (25a)$$

$$= (1 - r)\dot{\Sigma}_{\star,0}(R_{\text{gal}}) \int_0^T f(t|R_{\text{gal}}) dt \quad (25b)$$

$$\Rightarrow \dot{\Sigma}_{\star,0}(R_{\text{gal}}) = \Sigma_{\star}(R_{\text{gal}}) \left[(1 - r) \int_0^T f(t|R_{\text{gal}}) dt \right]^{-1} \quad (25c)$$

$$= \Sigma_{\star,0}g(R_{\text{gal}}) \left[(1 - r) \int_0^T f(t|R_{\text{gal}}) dt \right]^{-1} \quad (25d)$$

- Integrate surface density over area of the disc and you get the present day Milky Way stellar mass. This solves for the unknown $\Sigma_{\star,0}$:

$$M_{\star}^{\text{MW}} = \int_0^R \Sigma_{\star}(R_{\text{gal}}) 2\pi R_{\text{gal}} dR_{\text{gal}} \quad (26a)$$

$$= \Sigma_{\star,0} \int_0^R g(R_{\text{gal}}) 2\pi R_{\text{gal}} dR_{\text{gal}} \quad (26b)$$

$$\Rightarrow \Sigma_{\star,0} = M_{\star}^{\text{MW}} \left[\int_0^R g(R_{\text{gal}}) 2\pi R_{\text{gal}} dR_{\text{gal}} \right]^{-1} \quad (26c)$$

- Combine the last two equations into $\dot{\Sigma}_{\star}(R_{\text{gal}}, t)$ and obtain the following equation:

$$\dot{\Sigma}_{\star}(R_{\text{gal}}, t) = A f(t|R_{\text{gal}}) g(R_{\text{gal}}) \quad (27)$$

where

$$A = M_{\star}^{\text{MW}} \left[(1 - r) \int_0^R g(R_{\text{gal}}) 2\pi R_{\text{gal}} dR_{\text{gal}} \int_0^T f(t|R_{\text{gal}}) dt \right]^{-1} \quad (28)$$

This result makes intuitive sense: $f(t|R_{\text{gal}})$ specifies the time-dependence of the SFH and $g(R_{\text{gal}})$ specifies the radial dependence by construction, and M_{\star}^{MW} sets the overall normalization.

- This recipe implicitly assumes that radial migration does not significantly alter the surface density profile, and we have demonstrated in § 2.5 that this is the case for the Galactocentric radii of interest in this paper. It introduces scatter, but does not alter the overall dependence. This recipe can be employed in disc galaxy models as long as this is not violated.

B VARIATIONS IN STAR FORMATION EFFICIENCY

I'd be good to present a comparison of our fiducial, inside-out SFH model to a case where we make a significantly different assumption about τ_{\star} , and demonstrate that it doesn't impact the conclusions. This could motivate some of the discussion below as well, so that this Appendix can still tell a generally interesting story beyond just a comparison.

- Here are some of my thoughts on what could be causing the discrepancy with this τ_{\star} business. Since this is for discussion purposes, I haven't looked in detail at any of this; I'm simply speculating for the sake of discussion in the text.

• As discussed in § 2.6, our fiducial inside-out SFH overpredicts the molecular fraction f_{mol} relative to observed results in the Milky Way. This is a result of both an underpredicted HI + HII mass and an overpredicted H₂ mass. The predicted HI + HII mass is $5.54 \times 10^9 \text{ M}_{\odot}$ relative to an observed value of $\sim 10^{10} \text{ M}_{\odot}$ (Kalberla & Kerp 2009). The predicted H₂ mass is $7.14 \times 10^9 \text{ M}_{\odot}$, compared to an observed values of $2.5 \times 10^9 \text{ M}_{\odot}$ from Kalberla & Kerp (2009) and $1 \times 10^9 \text{ M}_{\odot}$ from Heyer & Dame (2015). Here we discuss potential sources of this discrepancy, and demonstrate that the assumed form of the star formation law does not impact our conclusions.

• First, the overpredicted H₂ mass. Our fiducial inside-out model, whose SFH is normalized to predict the same total stellar mass as that reported by Licquia & Newman (2015) with the surface density gradient of Bland-Hawthorn & Gerhard (2016) (see discussion in § 2.4 and 2.5). This model predicts a present-day SFR of $4.89 \text{ M}_{\odot} \text{ yr}^{-1}$, an overprediction relative to the observed value of $1.65 \text{ M}_{\odot} \text{ yr}^{-1}$ (Licquia & Newman 2015). However, our model predicts a gradient in the present-day surface density of star formation with radius which decreases monotonically (see the top row, left-center panel of Fig. 4), whereas the observations of Peek (2009) and Fraternali & Tomasetti (2012) would suggest that the Milky Way has begun the process of inside-out quenching. Fig. 2 of Fraternali & Tomasetti (2012) suggests that the surface density of star formation is significantly lower within $R_{\text{gal}} \lesssim 4 \text{ kpc}$. Cutting off our summation of the H₂ mass and present-day SFR at 4 kpc yields a present-day star formation rate of $3.06 \text{ M}_{\odot} \text{ yr}^{-1}$, and an H₂ mass of $4.48 \times 10^9 \text{ M}_{\odot}$, still an overprediction relative to Licquia & Newman (2015) and Kalberla & Kerp (2009). Instead cutting the summation off at 6 kpc yields $\dot{M}_{\star} = 1.97 \text{ M}_{\odot} \text{ yr}^{-1}$ and $\text{H}_2 = 2.88 \times 10^9 \text{ M}_{\odot}$, still a slight overprediction, but in better agreement still. Based on these calculations, we argue that the overpredicted H₂ mass in our fiducial inside-out model is a consequence of an overpredicted SFR at the present day. A model with either a more steeply declining SFH or quenching of star formation in the inner Galaxy as suggested by Peek (2009) and Fraternali & Tomasetti (2012) should resolve this discrepancy.

- Next, the under-predicted HI + HII mass. We speculate over a couple possible explanations of this.

– At first glance, this could point to a breakdown in the fundamental assumption of our adopted model for star formation efficiency. As discussed in § 2.6, we adopt the assumption that our model Galaxy follows the observed trend in the mean $\dot{\Sigma}_{\star} - \Sigma_{\text{gas}}$ relation, even though there is evidence that individual galaxies do not follow such a relation (Ellison et al. 2020a). Kennicutt & de los Reyes (2020) however argue that such a prescription should still be a valid recipe in galaxy evolution models.

– An important caveat in understanding our ISM masses is that because we treat the ISM as being available for star formation

in its entirety, they therefore track only the amount of gas in the *star forming reservoir*. Zero metallicity gas infall simply appears in each annulus to balance the amount of gas lost to star formation and outflows at each timestep in order to sustain the level of star formation at the next timestep. Based on our model, if gas is removed from the star forming reservoir, but remains in a spatially confined region reasonably defining the Galaxy, our model will treat it as an outflow and remove it from the ISM, though in nature it would still show up in observations and be counted as part of the Milky Way’s gas content. It’s therefore plausible that the difference could simply be nuanced semantics regarding what we refer to as an outflow numerically and what would be considered an outflow in an observational study (i.e. ejection from the disc).

- Based on this argument, one potential source of the difference could be infall from the circum- and inter-galactic media that have not yet been made available for star formation. This would show up as HI and HII observed in the ISM, but would not be counted in our models if it is not yet incorporated into the star forming reservoir.

- Another potential source could be star formation efficiency in the meaning that the ISM/feedback literature often takes: the fraction of molecular cloud masses that are converted into stars. To first order, this value is a few percent with some intrinsic scatter ([citation](#)). This means that in our models, a significant multiplier (~30) of the stellar mass formed in a given timestep should be added to the hot phase of the ISM in the form of HII through feedback processes. If this gas is never again made available for star formation, our model would treat it as an outflow, removing it from the ISM mass, while in nature it is still observed within the reasonably confined region defining the Galaxy.

- Together, these effects may constitute the remaining factor of \sim of warm and hot gas that our fiducial model is missing. However, an investigation of the impact of a multi-phase ISM on models such as these is outside the scope of this paper.

- Whatever the source of this discrepancy may be, the decision about the mathematical form of τ_\star does not significantly impact our conclusions, as demonstrated in the remainder of this Appendix.

REFERENCES

- Abolfathi B., et al., 2018, [ApJS](#), **235**, 42
 Adams S. M., Kochanek C. S., Gerke J. R., Stanek K. Z., Dai X., 2017, [MNRAS](#), **468**, 4968
 Ahumada R., et al., 2020, [ApJS](#), **249**, 3
 Andrews B. H., Weinberg D. H., Schönrich R., Johnson J. A., 2017, [ApJ](#), **835**, 224
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, [ARA&A](#), **47**, 481
 Astropy Collaboration et al., 2013, [A&A](#), **558**, A33
 Astropy Collaboration et al., 2018, [AJ](#), **156**, 123
 Basinger C. M., Kochanek C. S., Adams S. M., Dai X., Stanek K. Z., 2020, arXiv e-prints, p. [arXiv:2007.15658](#)
 Berg D. A., Skillman E. D., Croxall K. V., Pogge R. W., Moustakas J., Johnson-Groh M., 2015, [ApJ](#), **806**, 16
 Berg D. A., Pogge R. W., Skillman E. D., Croxall K. V., Moustakas J., Rogers N. S. J., Sun J., 2020, [ApJ](#), **893**, 96
 Bigiel F., Leroy A., Walter F., Blitz L., Brinks E., de Blok W. J. G., Madore B., 2010, [AJ](#), **140**, 1194
 Bird J. C., Loebman S. R., Weinberg D. H., Brooks A., Quinn T. R., Christensen C. R., 2020, arXiv e-prints, p. [arXiv:2005.12948](#)
 Bland-Hawthorn J., Gerhard O., 2016, [ARA&A](#), **54**, 529
 Brooks A. M., Zolotov A., 2014, [ApJ](#), **786**, 87
 Brown J. S., et al., 2019, [MNRAS](#), **484**, 3785
 Buck T., 2020, [MNRAS](#), **491**, 5435
 Buck T., Obreja A., Macciò A. V., Minchev I., Dutton A. A., Ostriker J. P., 2020, [MNRAS](#), **491**, 3461
 Chiappini C., Matteucci F., Gratton R., 1997, [ApJ](#), **477**, 765
 Chiappini C., Matteucci F., Romano D., 2001, [ApJ](#), **554**, 1044
 Chiappini C., et al., 2015, [A&A](#), **576**, L12
 Chisholm J., Tremonti C., Leitherer C., 2018, [MNRAS](#), **481**, 1690
 Christensen C., Quinn T., Governato F., Stilp A., Shen S., Wadsley J., 2012, [MNRAS](#), **425**, 3058
 Clarke A. J., et al., 2019, [MNRAS](#), **484**, 3476
 Cristallo S., et al., 2011, [ApJS](#), **197**, 17
 Daflon S., Cunha K., de la Reza R., Holtzman J., Chiappini C., 2009, [AJ](#), **138**, 1577
 Dalcanton J. J., 2007, [ApJ](#), **658**, 941
 de los Reyes M. A. C., Kennicutt Robert C. J., 2019, [ApJ](#), **872**, 16
 Edvardsson B., Andersen J., Gustafsson B., Lambert D. L., Nissen P. E., Tomkin J., 1993, [A&A](#), **500**, 391
 Ellison S. L., Lin L., Thorp M. D., Pan H.-A., Scudder J. M., Sanchez S. F., Bluck A. F. L., Maiolino R., 2020a, arXiv e-prints, p. [arXiv:2012.04771](#)
 Ellison S. L., Lin L., Thorp M. D., Pan H.-A., Sanchez S. F., Bluck A. F. L., Belfiore F., 2020b, arXiv e-prints, p. [arXiv:2012.04772](#)
 Feuillet D. K., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), **477**, 2326
 Feuillet D. K., Frankel N., Lind K., Frinchaboy P. M., García-Hernández D. A., Lane R. R., Nitschelm C., Roman-Lopes A. r., 2019, [MNRAS](#), **489**, 1742
 Fraternali F., Tomasetti M., 2012, [MNRAS](#), **426**, 2166
 Frinchaboy P. M., et al., 2013, [ApJ](#), **777**, L1
 Gaia Collaboration et al., 2018, [A&A](#), **616**, A1
 García-Benito R., et al., 2017, [A&A](#), **608**, A27
 García Pérez A. E., et al., 2016, [AJ](#), **151**, 144
 Gerke J. R., Kochanek C. S., Stanek K. Z., 2015, [MNRAS](#), **450**, 3289
 González Delgado R. M., et al., 2014, [A&A](#), **562**, A47
 Grand R. J. J., et al., 2017, [MNRAS](#), **467**, 179
 Grand R. J. J., et al., 2018, [MNRAS](#), **474**, 3629
 Griffith E., et al., 2020, arXiv e-prints, p. [arXiv:2009.05063](#)
 Grisoni V., Spitoni E., Matteucci F., Recio-Blanco A., de Laverny P., Hayden M., Mikolaitis Š., Worley C. C., 2017, [MNRAS](#), **472**, 3637
 Harris C. R., et al., 2020, [Nature](#), 585, 357
 Hayden M. R., et al., 2014, [AJ](#), **147**, 116
 Hayden M. R., et al., 2015, [ApJ](#), **808**, 132
 Haywood M., 2008, [MNRAS](#), **388**, 1175
 Heyer M., Dame T. M., 2015, [ARA&A](#), **53**, 583
 Holmberg J., Nordström B., Andersen J., 2007, [A&A](#), **475**, 519

- Holoien T. W. S., et al., 2019, *MNRAS*, **484**, 1899
 Holtzman J. A., et al., 2015, *AJ*, **150**, 148
 Hunter J. D., 2007, *Computing in Science & Engineering*, **9**, 90
 Isern J., 2019, *ApJ*, **878**, L11
 Johnson J. W., Weinberg D. H., 2020, *MNRAS*, **498**, 1364
 Kalberla P. M. W., Kerp J., 2009, *ARA&A*, **47**, 27
 Kalirai J. S., Hansen B. M. S., Kelson D. D., Reitzen D. B., Rich R. M., Richer H. B., 2008, *ApJ*, **676**, 594
 Kennicutt Robert C. J., 1998, *ApJ*, **498**, 541
 Kennicutt R. C., Evans N. J., 2012, *ARA&A*, **50**, 531
 Kennicutt Robert C. J., de los Reyes M. A. C., 2020, arXiv e-prints, p. arXiv:2012.05363
 Kroupa P., 2001, *MNRAS*, **322**, 231
 Krumholz M. R., Burkhardt B., Forbes J. C., Crocker R. M., 2018, *MNRAS*, **477**, 2716
 Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, **136**, 2782
 Leroy A. K., et al., 2013, *AJ*, **146**, 19
 Licquia T. C., Newman J. A., 2015, *ApJ*, **806**, 96
 Liu L., Gao Y., Greve T. R., 2015, *ApJ*, **805**, 31
 Loebman S. R., Ivezić Ž., Quinn T. R., Governato F., Brooks A. M., Christensen C. R., Jurić M., 2012, *ApJ*, **758**, L23
 Loebman S. R., et al., 2014, *ApJ*, **794**, 151
 Majewski S. R., et al., 2017, *AJ*, **154**, 94
 Maoz D., Graur O., 2017, *ApJ*, **848**, 25
 Maoz D., Mannucci F., 2012, *Publ. Astron. Soc. Australia*, **29**, 447
 Martig M., et al., 2016, *MNRAS*, **456**, 3655
 Matteucci F., Francois P., 1989, *MNRAS*, **239**, 885
 Melioli C., Brighenti F., D'Ercole A., de Gouveia Dal Pino E. M., 2008, *MNRAS*, **388**, 573
 Melioli C., Brighenti F., D'Ercole A., de Gouveia Dal Pino E. M., 2009, *MNRAS*, **399**, 1089
 Minchev I., Chiappini C., Martig M., 2013, *A&A*, **558**, A9
 Minchev I., Chiappini C., Martig M., 2014, *A&A*, **572**, A92
 Minchev I., Steinmetz M., Chiappini C., Martig M., Anders F., Matijevic G., de Jong R. S., 2017, *ApJ*, **834**, 27
 Mor R., Robin A. C., Figueiras F., Roca-Fàbrega S., Luri X., 2019, *A&A*, **624**, L1
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, **490**, 493
 Nidever D. L., et al., 2014, *ApJ*, **796**, 38
 Noguchi M., 2018, *Nature*, **559**, 585
 Nordström B., Andersen J., Holmberg J., Jørgensen B. R., Mayor M., Pont F., 2004a, *Publ. Astron. Soc. Australia*, **21**, 129
 Nordström B., et al., 2004b, *A&A*, **418**, 989
 Pagel B. E. J., Patchett B. E., 1975, *MNRAS*, **172**, 13
 Peek J. E. G., 2009, *ApJ*, **698**, 1429
 Pejcha O., Thompson T. A., 2015, *ApJ*, **801**, 90
 Pinsonneault M. H., et al., 2014, *ApJS*, **215**, 19
 Recio-Blanco A., et al., 2014, *A&A*, **567**, A5
 Rojas-Arriagada A., et al., 2017, *A&A*, **601**, A140
 Romano D., Karakas A. I., Tosi M., Matteucci F., 2010, *A&A*, **522**, A32
 Roškar R., Debattista V. P., Stinson G. S., Quinn T. R., Kaufmann T., Wadsley J., 2008a, *ApJ*, **675**, L65
 Roškar R., Debattista V. P., Quinn T. R., Stinson G. S., Wadsley J., 2008b, *ApJ*, **684**, L79
 Salpeter E. E., 1955, *ApJ*, **121**, 161
 Sánchez S. F., 2020, *ARA&A*, **58**, 99
 Schönrich R., Binney J., 2009, *MNRAS*, **396**, 203
 Sellwood J. A., Binney J. J., 2002, *MNRAS*, **336**, 785
 Sharma S., Hayden M. R., Bland-Hawthorn J., 2020, arXiv e-prints, p. arXiv:2005.03646
 Silva Aguirre V., et al., 2018, *MNRAS*, **475**, 5487
 Spitoni E., Recchi S., Matteucci F., 2008, *A&A*, **484**, 743
 Spitoni E., Matteucci F., Recchi S., Cescutti G., Pipino A., 2009, *A&A*, **504**, 87
 Spitoni E., Vincenzo F., Matteucci F., Romano D., 2016, *MNRAS*, **458**, 2541