The Equilibrium Model of Radial Metallicity Gradients in Disk Galaxies: Why Old Stars and High Redshift Galaxies are Metal-Rich

JAMES W. JOHNSON [iD][1,2,3] AND COLLABORATORS

[1]*Carnegie Science Observatories, 813 Santa Barbara St., Pasadena, CA 91101, USA*
[2]*Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA*
[3]*Center for Cosmology & Astroparticle Physics (CCAPP), The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA*

ABSTRACT

Metallicities of both gas and stars decline toward large radii in spiral galaxies, a smooth transition known as the radial metallicity gradient. We quantify the evolution in this relation in the Milky Way as traced by APOGEE red giants with age estimates from machine learning algorithms, finding minimal evolution in slope and no evolution in normalization up to ages of ~9 Gyr. This result challenges current models of Galactic chemical evolution, which typically predict metal abundances to decline toward old ages across much of the Galaxy, because ongoing star formation facilitates ongoing metal production. We therefore present the equilibrium model, a generic class of chemical evolution models whose defining feature is that gas-phase metallicities asymptotically approach a steady state whose value declines with radius in a manner that tracks the observations. Wind driven outflows with a mass loading factor that increases exponentially with radius are one possible origin of this behavior. We demonstrate using a fiducial choice of parameters that the equilibrium model offers a qualitative explanation for stellar metallicities that are age-independent at fixed radius even for old populations. If extended to other spiral galaxies, the equilibrium model offers a qualitative explanation to the surprisingly high metallicities observed at high redshift with JWST. The equilibrium model also suggests that the metallicity gradient is indirectly as opposed to directly related to the inside-out formation of the Galactic disk. The decline in metallicity is instead predicted to follow a decline in the time-averaged ratio of star formation per unit accretion toward large radii.

Corresponding author: James W. Johnson
jjohnson10@carnegiescience.edu

1. INTRODUCTION

Spiral galaxies are metal-rich at small radii and metal-poor at large radii. The metal mass fraction Z in both gas (e.g. Wyse & Silk 1989; Zaritsky 1992) and stars (e.g. Chen et al. 2003; Daflon & Cunha 2004; Cheng et al. 2012) declines exponentially with Galactocentric radius, a smooth transition known as the radial metallicity gradient. Thanks to the success of spectroscopic surveys, measurements are now available for thousands of external galaxies out to increasingly high redshift (e.g. Maiolino & Mannucci 2019; Sánchez 2020). First quantified through nebular emission lines in HII regions (Aller 1942; Searle 1971; Shaver et al. 1983), their presence and ubiquity are generally interpreted as evidence of "inside-out" disk growth, in which the inner disk assembles first and the outskirts follow suit on longer timescales (e.g. Matteucci & Francois 1989; White & Frenk 1991; Kauffmann 1996; Bird et al. 2013).

Because stars inherit the chemical composition of the local interstellar medium (ISM) when they form (to within $\lesssim 0.02 - 0.03$ dex; De Silva et al. 2006; Bovy 2016; Liu et al. 2016; Casamiquela et al. 2020), a common approach to constrain the evolution in the disk metallicity gradient is to analyze mono-age stellar populations. Application of this procedure is limited to the Milky Way (MW) where large samples of resolved stars are feasible. Despite the number of observationally accessible targets, the evolution in the ISM

metallicity gradient is only weakly constrained. One source of uncertainty is radial migration (e.g. Sellwood & Binney 2002), which can carry stars several kpc from their birth radius where their abundances no longer reflect the Galactic region in which they formed. Another source is the difficulty associated with precision age measurements for stars (see e.g. the reviews by Soderblom 2010 and Chaplin & Miglio 2013). Many investigations thus far have used spectral types that coarsely trace underlying populations at young, intermediate, and old ages, such as OB stars (e.g. Daflon & Cunha 2004), Cepheid Variables (e.g. Andrievsky et al. 2004; Luck et al. 2006, 2011; Luck & Lambert 2011; Yong et al. 2006), open clusters (e.g. Friel 1995; Chen et al. 2003; Magrini et al. 2009), and planetary nebulae (e.g. Maciel et al. 2003; Henry et al. 2010; Stanghellini & Haywood 2010; Magrini et al. 2016).

Reliable age measurements for red giants are particularly valuable for constraining the enrichment history of the MW disk, because they are accessible at large distances due to their high luminosities. Of the current methods for measuring stellar ages, asteroseismology is the most reliable for red giants. Modeling their photometric variability constrains their masses and evolutionary states (e.g. De Ridder et al. 2009; Bedding et al. 2010, 2011; Hekker & Christensen-Dalsgaard 2017), which enables an age inference based on the mass-lifetime relation (e.g. Larson 1974; Maeder & Meynet 1989;

Padovani & Matteucci 1993; Kodama & Arimoto 1997; Hurley et al. 2000). Joint catalogs of seismic ages and metal abundances from spectroscopic surveys (e.g. APOKASC; Pinsonneault et al. 2014, 2018) are therefore highly valuable. Some authors have trained neural networks on these catalogs in order to predict the asteroseismic data from the spectrum and construct large samples of ages (e.g. Ness et al. 2016; Mackereth et al. 2019; Leung et al. 2023; Stone-Martinez et al. 2024).

The combination of asteroseismology and spectroscopy has been a popular choice in recent investigations of the age-abundance structure of the Galactic disk (e.g. Silva Aguirre et al. 2018; Spitoni et al. 2019; Lian et al. 2020a,b; Matsuno et al. 2021). A result of particular interest to this paper is that old and intermediate-aged stars are not significantly more metal poor than young stars. Willett et al. (2023) demonstrate that conventional models of Galactic chemical evolution (GCE) underpredict the metallicities of old populations. Gallart et al. (2024) showed that stellar metallicity does not correlate significantly with age up to ~10 Gyr in six different age catalogs (see their Fig. 13; Xiang & Rix 2022; Queiroz et al. 2023; Kordopatis et al. 2023). The relationship between metallicity and Galactic radius also does not correlate significantly with age in the latest catalogs of open clusters and Cepheid variables (Spina et al. 2022; da Silva et al. 2023). These results are likely related to the surprisingly high metallicities observed in high redshift galaxies with the James Webb Space Telescope (JWST; see discussion in section 5.6 below).

The apparent lack of decline in stellar metallicities across such a broad range of age demands a theoretical explanation. Such a result is not predicted by "classical" GCE models, which would postulate that metallicities should decline toward high ages since ongoing star formation facilitates ongoing metal production (see e.g. the reviews by Tinsley 1980 and Matteucci 2021). To that end, this paper proposes the *equilibrium model* of radial metallicity gradients in disk galaxies. Because we invoke this model in an interpretive manner, our discussion casts this equilibrium scenario as a class of models rather than one specific parameterization. The defining feature of the equilibrium model is that the ISM metallicity Z at fixed radius does not evolve significantly with time after the first ~few Gyr of disk evolution. Instead, Z evolves toward a local equilibrium abundance Z_{eq}, which exponentially with radius in a manner that tracks the observed gradient. The roots of equilibrium chemical evolution can be traced back to Larson (1974), who showed that in the presence of ongoing accretion, ISM abundances tend toward an equilibrium value (hence the model's namesake) at which metal production by stars is balanced by hydrogen gained through accretion.

Variations in the equilibrium abundance have played a central role in models of the mass-metallicity relation (MZR) for galaxies over the past ~15 years. High mass galaxies tend to be more metal-rich than their low mass counterparts in terms of both gas (e.g. Tremonti et al. 2004; Andrews & Martini 2013; Zahid et al. 2012; Blanc et al. 2019) and stellar populations (e.g. Gallazzi et al. 2005; Kirby et al. 2013; Simon 2019). This trend is often attributed to low mass galaxies more read-

ily ejecting gas in a wind due to their weak gravitational fields (e.g. Finlator & Davé 2008; Peeples & Shankar 2011; Lilly et al. 2013; Sanders et al. 2021; Chartab et al. 2023). While this literature invokes variations in chemical equilibrium *between* galaxies, our proposed equilibrium scenario invokes these variations *within* individual spiral galaxies.

In section 2 below, we present evidence in support of the argument that stellar metallicities are age-independent up to at least ~9 − 10 Gyr. We construct a set of GCE models demonstrating proof of concept of the equilibrium scenario in section 3. In section 4, we compare the predictions of these models with our sample and highlight the key differences between parameter choices. The equilibrium scenario has a number of implications in the broader context of galactic astrophysics, which we discuss individually in section 5. We summarize our findings in section 6.

2. EMPIRICAL AGE AND METALLICITY GRADIENTS

There are many spectroscopic surveys to choose from to characterize the age-abundance structure of the Galactic disk, such as LAMOST (Luo et al. 2015), GALAH (De Silva et al. 2015; Martell et al. 2017), *Gaia*-ESO (Gilmore et al. 2012), and APOGEE[1] (Majewski et al. 2017). APOGEE is particularly conducive to this task, because it targets luminous evolved stars accessible at large distances and is less susceptible to dust obscuration with spectra taken at near-IR wavelengths ($\lambda = 1.51 − 1.70$ μm; Wilson et al. 2019).

---

[1] LAMOST: Large sky Area Multi-Object fibre Spectroscopic Telescope

GALAH: GALactic Archaeology with Hermes

APOGEE: Apache Point Observatory Galaxy Evolution Experiment

2.1. *The Sample*

In this paper, we focus on the AstroNN value added catalog[2] for APOGEE's seventeenth data release (DR17; Abdurro'uf et al. 2022). In making the age measurements for the original value added catalog, Mackereth et al. (2019) used AstroNN (Leung & Bovy 2019a) to train a Bayesian convolutional neural network on APOGEE DR14 spectra and asteroseismic ages from APOKASC-2 (Pinsonneault et al. 2018). Retrained on DR17 spectra, the new catalog improves the performance at low metallicity by incorporating additional asteroseismic data from Montalbán et al. (2021) and provides individual stellar abundances (Leung & Bovy 2019a) and distances (Leung & Bovy 2019b) through *Gaia*-eDR3 (Gaia Collaboration et al. 2021). We discuss the differences between AstroNN and the recent Leung et al. (2023) catalog in section 2.4 below.

We filter the sample based on the following selection criteria:

- STAR_BAD == 0

- EXTRATARG == 0

- S/N ≥ 80

- $\log g = 1 − 3.8$

- $T_{\text{eff}} = 3400 − 5500$ K

To avoid potential contamination by the main sequence, we additionally exclude stars that have both $\log g > 3$ and $T_{\text{eff}} < 4000$ K. Age uncertainties in the AstroNN catalog

---

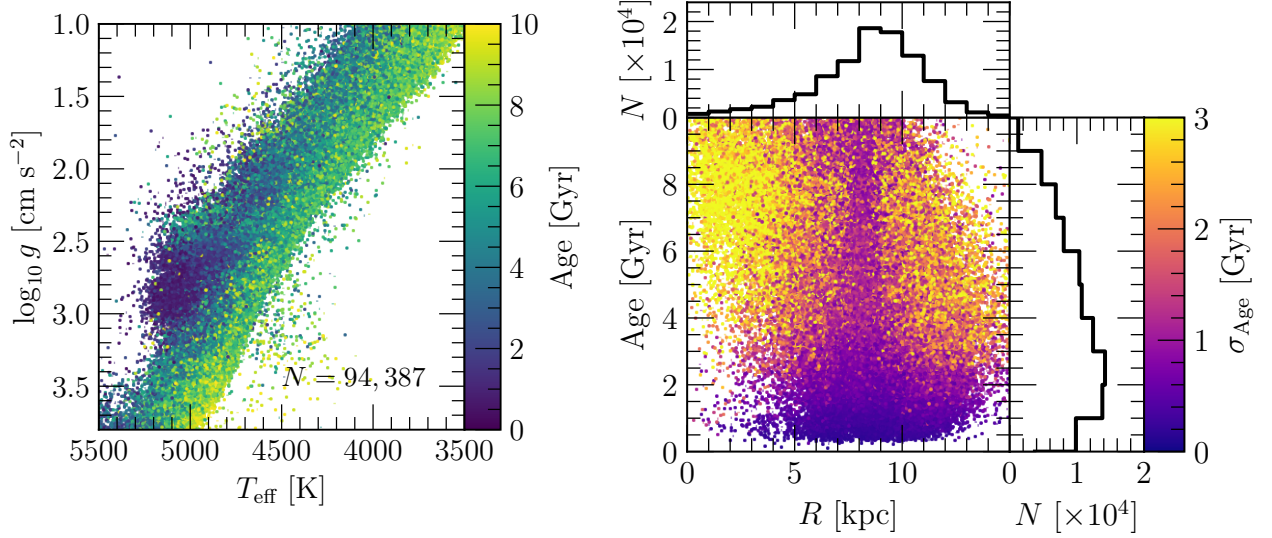[2] https://www.sdss.org/dr18/data_access/value-added-catalogs/?vac_id=85

Figure 1. Our sample: the AstroNN value added catalog (Mackereth et al. 2019) restricted to stars within $|z| \leq 0.5$ kpc of the disk midplane and the ranges of stellar parameters and Galactocentric radius visualized in these panels (see discussion in section 2.1). **Left**: The Kiel diagram, color coded by stellar age according to the colorbar. **Right**: The Galactocentric radii and ages of each star, color coded by the reported age uncertainties according to the colorbar. Top and right panels show distributions in radius and age, respectively. **Summary**: By drawing stars from across the red giant branch, we construct a large sample with excellent coverage of the Galactic disk (see discussion in section 2.1).

become substantial for stars older than $\tau \gtrsim 8 - 10$ Gyr (Leung et al. 2023). Since we are most interested in thin disk populations anyway, we impose additional cuts on age, radius, and midplane distance of $\tau \leq 10$ Gyr, $R \leq 15$ kpc, and $|z| \leq 0.5$ kpc, respectively. These criteria yield a final sample of $N = 94,387$ red giant and red clump stars.

The left panel of Fig. 1 shows the Kiel diagram of our sample color coded by stellar age. Young populations are preferentially located in the red clump, whereas the oldest stars distribute themselves more evenly along the red giant branch. This effect is primarily driven by actual changes in the distribution of stars along the giant branch as a function of population age (e.g. Girardi 2016). However, systematic uncertainties in APOGEE tend to present as spurious correlations with $T_{\rm eff}$ and $\log g$ (e.g. Jönsson et al. 2018; Eilers

et al. 2022). Our main conclusions in this paper are driven by variations in abundances between mono-age populations, which could be affected by systematics if $T_{\rm eff}$ or $\log g$ act as confounding variables in metallicity and age (see discussion in section 2.4 below).

The right panel of Fig. 1 shows the Galactocentric radii and ages of each star along with the associated 1-D distributions. Our sample achieves excellent radial coverage of the disk, particularly in the $R = 5 - 12$ kpc range, where 81,580 of the 94,387 stars in our full sample reside. Coverage is best in the solar annulus, with 30,201 stars found between $R = 7$ and 9 kpc. Unless otherwise noted, we focus on bins of Galactocentric radius and age that are 1 kpc and 1 Gyr wide throughout this paper. We therefore color code the right panel of Fig. 1 by the overall age uncertainty σ_τ as opposed to the
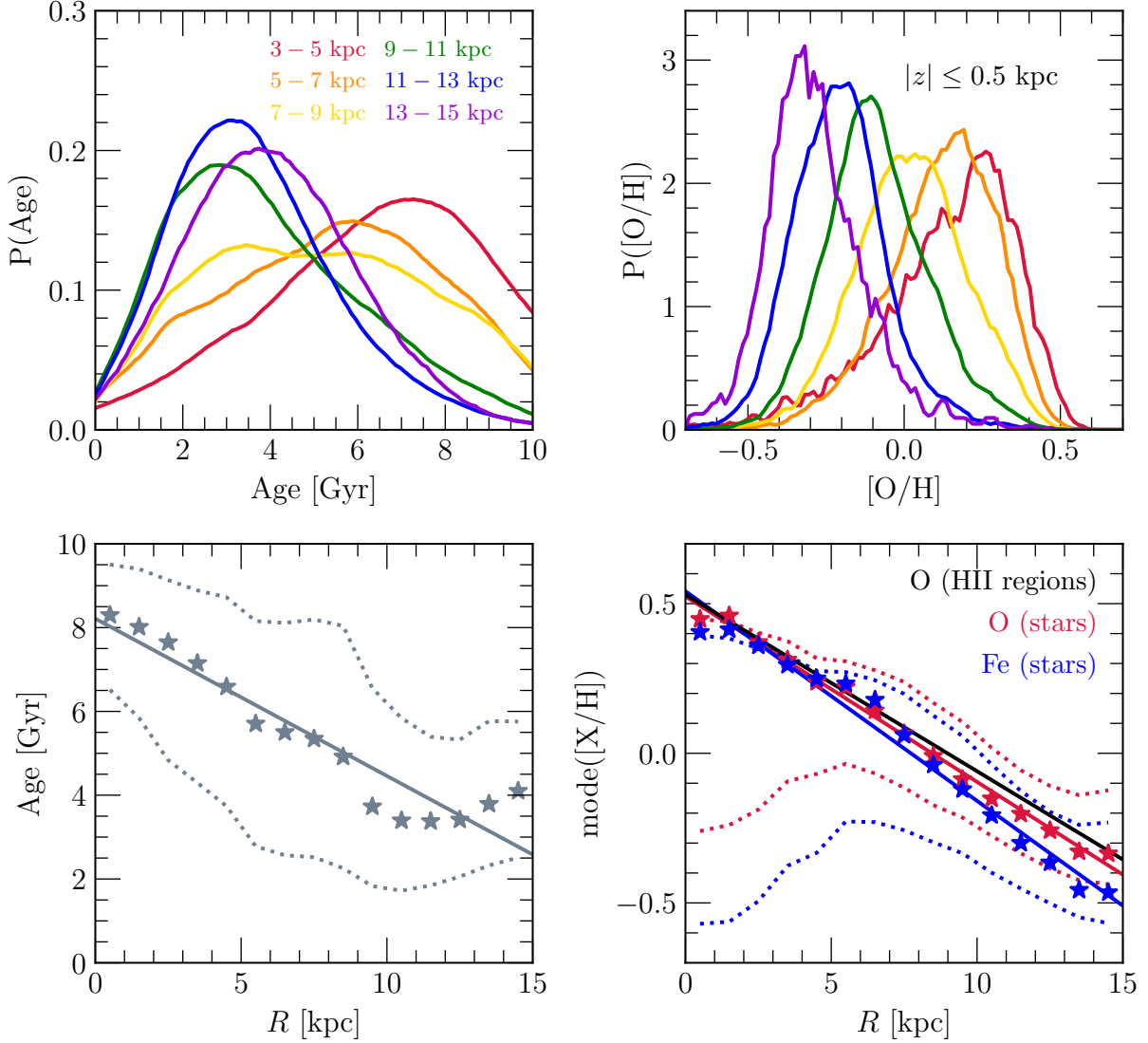
Figure 2. The age-abundance structure of the thin disk as traced by our sample (see Fig. 1 and discussion in section 2.1). **Top**: Distributions in stellar age (left) and [O/H] (right) in 2-kpc wide bins of Galactocentric radius, color coded according to the legend in the top-left panel. Each distribution is box-car smoothed with a window width of the median measurement uncertainty in each radial range. **Bottom**: Median age (left) and the peak of the metallicity distribution (right) in 1-kpc bins of radius (stars). Dotted lines denote the 16th and 84th percentiles of the distribution in each radial bin, while solid lines mark the line of best fit to the corresponding summary statistics (parameters of which are given in Table 1). We additionally plot the gradient in [O/H] measured from Galactic HII regions by Méndez-Delgado et al. (2022, black). **Summary**: In agreement with previous work, stellar populations tend to be younger and more metal-poor at large radii. The stellar and ISM metallicity gradients in the Galactic disk are consistent to within 1σ.

fractional uncertainty σ_τ/τ in order to illustrate the extent to which neighboring age bins are reliably distinct from one another.

The reported age uncertainties reach $\sigma_\tau \approx 1$ Gyr or better across much of the Galactic disk for young populations ($\tau \lesssim 3 - 4$ Gyr) and up to $\tau \approx 10$ Gyr near the Sun. Stellar abundances, however, have median uncertainties of $\sigma_{[Fe/H]} = 0.0087$ and $\sigma_{[O/Fe]} = 0.017$, which is sufficiently precise for our purposes. We discuss sources of uncertainty further in section 2.4 below.

2.2. *Radial Gradients*

The top panels of Fig. 2 show age and abundance distributions in our sample in 2-kpc bins of Galactocentric radius. We have not made any corrections for the survey selection function, so these are distributions of observed APOGEE stars satisfying our selection cuts as opposed to mass-weighted distributions. In agreement with previous work (see discussion in section 1), stars tend to be young in the outer Galaxy and old in the inner Galaxy, with tails toward old and young populations, respectively. Abundances follow a similar pattern, shifting from a metal-rich mode to a metal-poor mode with increasing radius.

We quantify the strength of these radial gradients by computing summary statistics of each distribution in 1-kpc bins of Galactocentric radius. The bottom left panel of Fig. 2 shows the median age and the 16th and 84th percentiles of the age distribution as a function of radius. A linear regression indicates a median trend of $\nabla\tau_{1/2} = -0.375 \pm 0.036$ Gyr/kpc with an intercept of 8.21 ± 0.31 Gyr. The observed median

age closely follows this line of best-fit within $R \lesssim 9$ kpc, beyond which ages begin to increase slightly with radius.

While we quantify the radial age gradient in terms of a median trend, we use the mode to quantify the metallicity gradient. We discuss our motivation behind this choice in section 5.2 below. In short, our GCE models suggest that the mode is less susceptible to modification by stellar migration than the mean and median. As a result, the peak of the metallicity distribution function (MDF) is a better proxy for the ISM abundance at a given radius and lookback time. To mitigate noise in the inferred mode introduced by poisson fluctuations, we first fit a skew normal distribution to the MDF in each radial bin and determine the positions of the peaks with optimization.

The top right panel of Fig. 2 shows the resultant gradients in [O/H] and [Fe/H].[3] As expected, stars tend to decline in metallicity with increasing radius. Linear regressions indicate slopes of $\nabla[O/H] = -0.062 \pm 0.001$ kpc^{-1} and $\nabla[Fe/H] = -0.070 \pm 0.003$ kpc^{-1}, in reasonable agreement with previous measurements from APOGEE (e.g. Frinchaboy et al. 2013; Myers et al. 2022). For comparison, we additionally plot Méndez-Delgado et al.'s (2022) fit to the gas-phase O gradient traced by Galactic HII regions including their estimated correction for temperature inhomogeneities. The two O gradients are consistent within their 1σ uncertainties.

---

[3] We follow conventional notation where $[X/Y] \equiv \log_{10}(N_X/N_Y) - \log_{10}(N_X/N_Y)_\odot$.

2.3. *Evolution of the Abundance Gradient*

In this section, we quantify the evolution of the disk abundance structure by repeating our measurements in section 2.2 above in 1 Gyr bins of stellar age. After sorting based on both age and radius, we fit for the mode of the MDF only if the bin contains at least 200 stars. We estimate statistical uncertainties using jackknife resampling.

The top panels of Fig. 3 show the resulting relationship between metallicity and Galactocentric radius for mono-age populations. While there are some noticeable variations in the slope, the result that is of particular importance to this paper is that the *normalization* of this relationship does not decline significantly with age across a broad range. To demonstrate this point further, the lower left panel shows the [O/H] distributions in the solar annulus ($R = 7 - 9$ kpc) in the same 1-Gyr age bins. There is little to no variation in the centroid with age. If anything, the tail of the distribution shifts toward slightly super-solar abundances for old populations, which would only strengthen the argument that old stars are more metal-rich than predicted by conventional GCE models (see discussion in section 1).

We apply linear regressions to these [O/H] $- R$ and [Fe/H] $- R$ trends. The lower middle and lower right panels of Fig. 3 show the slopes and values at $R = 8$ kpc as functions of the age bin, which we also report in Table 1. Our measurements indicate that, to first order, both the normalization and slope were established $\sim 8 - 9$ Gyr ago. For comparison, the error bar in the lower-right panel shows [O/H] $= 0 \pm 0.05$, which corresponds to the level of non-axisymmetric varia-

Table 1. A summary of the linear regressions applied to radial metallicity and age gradients in this paper (see discussion in sections 2 and 5.3). We use a pivot point at $R_{gal} = 8$ kpc in all regressions (i.e. $y = m(R - R_\odot) + b$).

| Age Range	Slope	Value at 8 kpc
	[O/H]	
All Stars	-0.062 ± 0.001 kpc^{-1}	0.028 ± 0.006
≤ 9 Gyr	-0.062 ± 0.002 kpc^{-1}	0.028 ± 0.007
$0 - 1$ Gyr	-0.055 ± 0.011 kpc^{-1}	0.018 ± 0.022
$1 - 2$ Gyr	-0.056 ± 0.005 kpc^{-1}	0.028 ± 0.013
$2 - 3$ Gyr	-0.059 ± 0.005 kpc^{-1}	0.026 ± 0.016
$3 - 4$ Gyr	-0.066 ± 0.004 kpc^{-1}	0.014 ± 0.012
$4 - 5$ Gyr	-0.079 ± 0.004 kpc^{-1}	0.030 ± 0.012
$5 - 6$ Gyr	-0.080 ± 0.007 kpc^{-1}	0.051 ± 0.021
$6 - 7$ Gyr	-0.055 ± 0.008 kpc^{-1}	0.073 ± 0.025
$7 - 8$ Gyr	-0.049 ± 0.002 kpc^{-1}	0.125 ± 0.007
$8 - 9$ Gyr	-0.049 ± 0.007 kpc^{-1}	0.103 ± 0.030
$9 - 10$ Gyr	-0.001 ± 0.005 kpc^{-1}	0.029 ± 0.018
	[Fe/H]	
All Stars	-0.070 ± 0.003 kpc^{-1}	-0.019 ± 0.013
≤ 9 Gyr	-0.068 ± 0.003 kpc^{-1}	-0.023 ± 0.014
$0 - 1$ Gyr	-0.068 ± 0.010 kpc^{-1}	0.045 ± 0.020
$1 - 2$ Gyr	-0.072 ± 0.005 kpc^{-1}	0.028 ± 0.014
$2 - 3$ Gyr	-0.077 ± 0.006 kpc^{-1}	-0.006 ± 0.017
$3 - 4$ Gyr	-0.083 ± 0.005 kpc^{-1}	-0.044 ± 0.015
$4 - 5$ Gyr	-0.096 ± 0.007 kpc^{-1}	-0.032 ± 0.020
$5 - 6$ Gyr	-0.097 ± 0.012 kpc^{-1}	-0.028 ± 0.035
$6 - 7$ Gyr	-0.066 ± 0.012 kpc^{-1}	0.012 ± 0.036
$7 - 8$ Gyr	-0.051 ± 0.004 kpc^{-1}	0.082 ± 0.014
$8 - 9$ Gyr	-0.061 ± 0.008 kpc^{-1}	0.013 ± 0.031
$9 - 10$ Gyr	0.038 ± 0.006 kpc^{-1}	-0.207 ± 0.022
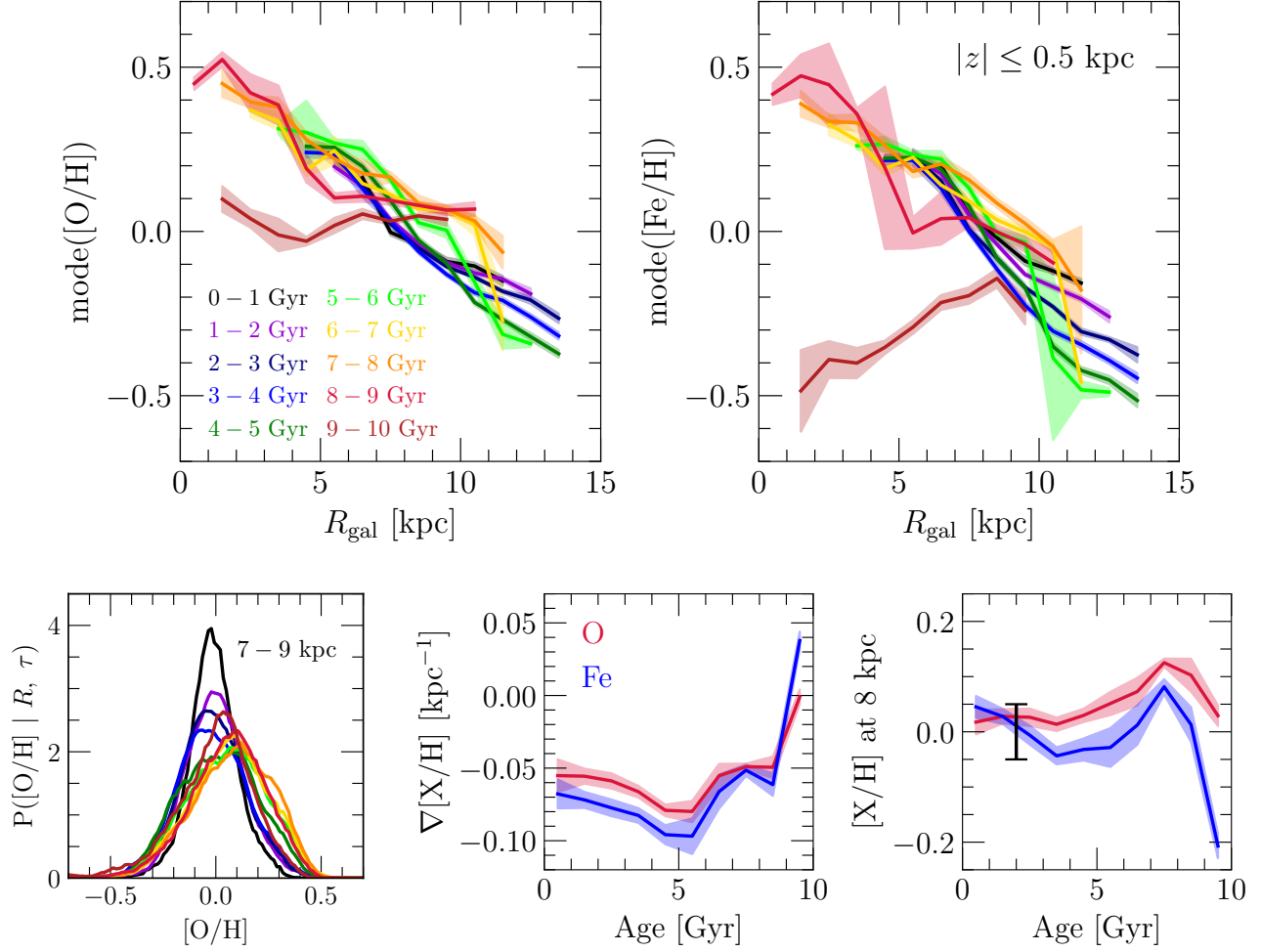
Figure 3. Radial metallicity gradients as a function of stellar age for stars within $|z| \leq 0.5$ kpc of the disk midplane. **Top**: The peak of the [O/H] (left) and [Fe/H] (right) distributions in 1-kpc bins of radius, color-coded by stellar age according to the legend in the top-left panel. Shaded regions denote the statistical uncertainty in the mode of the MDF estimated through jackknife resampling (see discussion in section 2.3). **Bottom**: The distribution in [O/H] in the solar annulus ($R = 7 - 9$ kpc; left) in the same age bins as the top panels. Middle and right panels show the slope and value at $R = 8$ kpc, respectively, inferred from linear regressions to the metallicity gradients in each age bin, parameters of which are reported in Table 1. The error bar in the right panel marks 0 ± 0.05, broadly consistent with Wenger et al.'s (2019) measurement of the non-axisymmetric variations in Galactic HII regions. **Summary**: The radial metallicity gradient of stars in the Galactic disk is largely age-independent up to ~9 Gyr in terms of both slope and normalization.

tions in ISM abundances predicted by simulations (Grand et al. 2016) and observed in HII regions (Wenger et al. 2019). The variations in stellar metallicities up to ages of ~8 − 10 Gyr are comparable to this intrinsic dispersion in ISM abundances. This apparent lack of change in stellar metallicities

with age, even for old populations, is our primary motivation in constructing the equilibrium model of metallicity gradients.

The gradient slope is less age independent than the normalization, but variations are still small in scale. It first held

steady at ∇[O/H] $\approx \nabla$[Fe/H] ≈ -0.05 kpc^{-1} for ~3 Gyr, then steepened to ∇[O/H] ≈ -0.07 kpc^{-1} and ∇[Fe/H] $= -0.09$ kpc^{-1} before gradually returning to its original value until the present day. Inspection of the top panels of Fig. 3 indicates that this steepening is likely driven by a brief decline in stellar metallicities in the outer disk ~6 Gyr ago. These fluctuations are small and do not significantly alter the structure of the relation between stellar metallicity and Galactic radius. In sections 4 and 4.2 below, we argue that the scale of these variations is accurately described by a perturbation of some equilibrium state by a merger event. Increases in the metal abundance with time across all radii predict larger differences between mono-age populations.

2.4. *Sources of Uncertainty*

Because the AstroNN ages were inferred by modeling the APOGEE spectra (see discussion in section 2.1), the values may be biased by correlations between stellar age and metal abundances (e.g. the age-[O/Fe] relation; Feuillet et al. 2018, 2019). To assess the impact of this potential systematic uncertainty, we replicate our measurements in Fig. 3 using the Leung et al. (2023) ages in Appendix A. They demonstrate that their ages are not sensitive to alpha and iron-peak element abundances, making their catalog a useful benchmark for verification. We find similar results with both catalogs, suggesting that these biases in AstroNN are not at a level that would alter our main conclusions. Although the Leung et al. (2023) ages are likely more reliable (see discussion in Appendix A), this catalog is a factor of ~2.5 smaller due to the narrower range of surface gravities in the training set. We

therefore focus on the AstroNN ages in this paper, which is nonetheless reliable in the age range we are most interested in anyway ($\tau \lesssim 8 - 10$ Gyr; see Fig. 11 of Leung et al. 2023 and discussion in their section 7.1).

Stellar ages themselves, including those that are not based on neural networks, are also a significant source of uncertainty (e.g. Soderblom 2010; Chaplin & Miglio 2013). If the uncertainties are particularly large, then the apparent lack of relationship with age seen in Fig. 3 could arise because different age bins do not actually trace distinct populations. However, in order for uncertainties to drive this conclusion, the measurements must be sufficiently imprecise to not tell the difference between ~1 − 3 and ~7 − 9 Gyr old stars. With or without machine learning, asteroseismology can at least reliably distinguish between young, intermediate, and old populations (e.g. Leung et al. 2023; Stone-Martinez et al. 2024). The result that metallicity does not decline substantially with age across this broad range should therefore be statistically significant even with the uncertainties involved. Our confidence in this conclusion is also supported by related investigations using different methods for determining ages, which have found similar results (Spina et al. 2022; da Silva et al. 2023; Gallart et al. 2024; see discussion in section 1).

Although we have not accounted for target selection in APOGEE (which is described in detail in Zasowski et al. 2013, 2017, Beaton et al. 2021, and Santana et al. 2021), these effects also should not affect our main empirical result. At a given distance, selection effects are introduced into our sample by the metallicity dependence of red giant lifetimes

and luminosities, which is minimal at disk-like abundances (e.g. Hurley et al. 2000). A representation such as Fig. 3 should therefore be relatively robust. In support of this argument, we note that Imig et al. (2023) did account for selection effects. Their measurements in mono-age populations also do not show an obvious decline in metallicity toward old ages across most of the Galactic disk (see their Fig. 17).

3. GALACTIC CHEMICAL EVOLUTION MODELS

Our GCE models are adapted from J21, which we integrate using the publicly available VERSATILE INTEGRATOR FOR CHEMICAL EVOLUTION[4] (VICE; Johnson & Weinberg 2020). Following previous models with similar motivations (e.g. Schönrich & Binney 2009; Minchev et al. 2013, 2014), these models discretize the Galactic disk into $\delta R = 100$ pc annuli from $R = 0$ to 20 kpc. By predicting abundances for multiple regions simultaneously, these so-called "multizone" models are a more realistic description of the Galactic disk than conventional one-zone GCE models (see e.g. the reviews by Tinsley 1980 and Matteucci 2021) and are significantly less computationally expensive than hydrodynamic simulations. A full detailed description of the framework can be found in section 2 of J21, with a shorter summary in section 3 of Johnson et al. (2023b). In this section, we review the model components relevant to this paper. Table 2 summarizes the key parameters.

---

[4] Install: https://pypi.org/project/vice

Documentation: https://vice-astro.readthedocs.io/en/latest

Source code: https://github.com/giganano/VICE.git

These models follow N,NNN,NNN stellar populations that form over a disk lifetime of $\tau_{\mathrm{disk}} = 13.2$ Gyr. We are motivated not by the precision of inferred parameter values but by proof of concept of the equilibrium scenario. We therefore refrain from detailed fits to the disk age-abundance structure observed by APOGEE, instead sticking to a handful of illustrative cases. The models we construct are somewhat idealized, treating the assembly history of the MW as one continuous episode of star formation. This choice allows us to highlight the defining characteristics of equilibrium chemical evolution in disk galaxies.

Following our measurements in section 2, we focus on alpha and iron-peak elements, taking O and Fe as representative cases thereof. Production of these metals is dominated by core collapse supernovae (CCSNe) and Type Ia supernovae (SNe Ia; Johnson 2019). Our yields are defined as the net mass production of either element in units of the mass of the progenitor stellar population. For example, a value of $y_x = 0.001$ would imply that a hypothetical 1000 $M_\odot$ star cluster would produce 1 $M_\odot$ of some element x. In the case of CCSNe, new metals are ejected to the ISM instantaneously, so the rate of change in the surface density of x follows according to

$$\dot{\Sigma}_x^{\mathrm{CC}} = y_x^{\mathrm{CC}} \dot{\Sigma}_\star,$$ (1)

where y_x^{CC} is the population-averaged yield from massive stars and $\dot{\Sigma}_\star$ is the local SFR. In the case of SNe Ia, production is spread out over the course of the delay-time distribution

Table 2. A summary of our GCE models (see discussion in section 3 for details).

Model Name	$\eta_\odot$	R_η	$y/Z_\odot$	Star formation History	Calibrated to Reproduce
Primary Set					
ηExp-y2	1.4	7 kpc	2	$f_{\text{rise-fall}}(t)$	$\tau_{1/2}(R)$ and $[\text{O/H}]_{\text{ISM}}(R)$
ηExp-y1	0.4	7 kpc	1	$f_{\text{rise-fall}}(t)$	$\tau_{1/2}(R)$ and $[\text{O/H}]_{\text{ISM}}(R)$
η0.4-y1	0.4	∞ (η = constant)	1	$f_{\text{rise-fall}}(t)$	$[\text{O/H}]_{\text{ISM}}(R)$
η0-y1	0	∞ (η = constant)	1	$f_{\text{rise-fall}}(t)$	$[\text{O/H}]_{\text{ISM}}(R)$
Variations of the ηExp-y2 model					
ηExp-y2-steep	1.4	4.3 kpc ($\nabla_{\text{eq}} = -0.04\ \text{kpc}^{-1}$)	2	$f_{\text{rise-fall}}(t)$	N/A (built upon ηExp-y3)
ηExp-y2-shallow	1.4	10.9 kpc ($\nabla_{\text{eq}} = -0.1\ \text{kpc}^{-1}$)	2	$f_{\text{rise-fall}}(t)$	N/A (built upon ηExp-y3)
ηevolExp-y2	1.4	7 kpc ($t \le 8$ Gyr), 21.7 kpc ($t > 8$ Gyr)	2	$f_{\text{rise-fall}}(t)$	N/A (built upon ηExp-y3)
ηExp-y2+burst	1.4	7 kpc	2	$f_{\text{rise-fall}}(t)\,(1 + A_b\Phi(t|t_b, \sigma_b, \alpha_b))$	N/A (built upon ηExp-y3)

Table 3. A summary of the stellar yields in our GCE models (see Table 2 and discussion in section 3 for details).

Yield	$y/Z_\odot = 1$	$y/Z_\odot = 2$
y_O^{CC}	0.0057	0.0114
y_O^{Ia}	0	0
y_{Fe}^{CC}	4.5×10^{-4}	9.0×10^{-4}
y_{Fe}^{Ia}	8.4×10^{-4}	0.00168

(DTD) R_{Ia} according to

$$\dot{\Sigma}_x^{\text{Ia}} = y_x^{\text{Ia}} \frac{\int_0^t \dot{\Sigma}_\star(t') R_{\text{Ia}}(t - t')dt'}{\int_0^\infty R_{\text{Ia}}(t')dt'}, \qquad (2)$$

where y_x^{Ia} is the population-averaged yield from SNe Ia. This quantity can be expressed as the product of the mean mass produced by a single SN Ia event and the mean number of events per unit mass of star formation.

Table 3 presents our adopted yield values. With these choices, 35% (65%) of Fe arises from CCSNe (SNe Ia), nearly the same breakdown as J21. Each of our GCE models follow one of two overall normalizations but with the same ratios of yields. We refer to these two choices collectively as $y/Z_\odot = 1$ and $y/Z_\odot = 2$, because the total yields of O and Fe are either equal to their solar abundances[5] or a factor of two higher. We show in sections 4 and 5.5 below that the overall scale of stellar yields is related to the timescale on which the disk reaches the equilibrium state.

These choices of stellar yields are based on a mix of theoretical and empirical considerations. In particular, Griffith et al. (2021) demonstrate that plausible variations in the amount of black hole formation can account for a factor of ~3 difference

---

[5] We use the measurements of O and Fe in the solar photosphere from Asplund et al. (2009).

in alpha element production. If most massive stars explode as a CCSN, then the predictions of typical models (e.g. Chieffi & Limongi 2013) are broadly consistent with the $y/Z_\odot = 3$ scenario. Weinberg et al. (2023) instead advocate for the lower $y/Z_\odot = 1$ scale based on Rodríguez et al.'s (2022) measurements of the mean ^{56}Ni yield from Type II SNe using the radioactive tails of their lightcurves. We discuss the normalization of stellar yields further in section 5.5 below.

Following J21, our models assume a fiducial choice of SFH given by

$$\dot{\Sigma}_\star \propto f_{\text{rise-fall}}(t) = \left(1 - e^{-t/\tau_{\text{rise}}}\right) e^{-t/\tau_{\text{sfh}}}, \qquad (3)$$

where τ_{rise} and τ_{sfh} control the timescales on which the SFR rises at early times and falls at late times, respectively. Their values differ between models (see discussion in section 3.2 and Appendix B). This prescription allows more control over the detailed shape of the SFH than a linear-exponential $te^{-t/\tau_{\text{sfh}}}$ form at the expense of an additional free parameter. The normalization of the SFH is set at each radius such that the predicted stellar mass and surface density gradient at the present day are consistent with the findings of Licquia & Newman (2015) and Bland-Hawthorn & Gerhard (2016), respectively. In section 3.3 below, we describe a model in which we impose an accretion-induced burst of star formation atop this underlying form.

Previous iterations of the J21 GCE models have implemented the radial migration of stars (e.g. Sellwood & Binney 2002) by "tagging" stellar populations with star particles from the h277 hydrodynamic simulation that formed at similar radii and times. Such an approach can be understood as enforcing

the dynamical history of h277 on the GCE model. Here, we use the updated version presented in Dubay et al. (2024; see their Appendix C). Present-day Galactocentric radii are determined by sampling from a normal distribution centered on the birth radius of a stellar population, while mid-plane distances $|z|$ are determined by sampling from a sech^2 function (Spitzer 1942). This approach closely approximates the distributions of final radius and mid-plane distance seen in h277 for mono-age populations formed in different Galactic regions. The advantage over the original tagging approach is that stellar populations born in the outer disk at early times were subject to sampling noise due to the rarity of these populations in the simulation.

We have also updated the star formation law used in these GCE models. In J21, we used a three-component power-law relationship between the surface densities of gas Σ_g and star formation $\dot{\Sigma}_\star$ based on observations by Bigiel et al. (2010) and Leroy et al. (2013; see Fig. 2 of Krumholz et al. 2018). The three component power-law complicates our parameter calibration, which determines τ_{rise} and τ_{sfh} as a function of Galactocentric radius in each of our GCE models (see discussion in section 3.2 and Appendix B). We therefore use a more classical single power-law prescription $\dot{\Sigma}_\star \propto \Sigma_g^N$ with $N = 1.5$ based on Kennicutt (1998) with a transition to a linear star formation law above $\Sigma_g = 10^8 M_\odot$ kpc^{-2}.

In detail, this relationship is implemented by computing the star formation efficiency (SFE) timescale $\tau_\star \equiv \Sigma_g/\dot{\Sigma}_\star$ (referred to as the "depletion time" by some authors) accord-

ing to

$$
\tau_\star =
\begin{cases}
\tau_{\mathrm{mol}} & (\Sigma_g \geq 10^8 M_\odot \mathrm{kpc}^{-2}) \\[2ex]
\tau_{\mathrm{mol}} \left(\dfrac{\Sigma_g}{10^8 \mathrm{kpc}^{-2}} \right)^{-1/2} & (\Sigma_g < 10^8 M_\odot \mathrm{kpc}^{-2}),
\end{cases}
\tag{4}
$$

where τ_{mol} is the value of $\tau_\star$ when all of the hydrogen is in the molecular state. We retain the prescription for τ_{mol} from J21, which takes $\tau_{\mathrm{mol}} = 2$ Gyr at the present day (Leroy et al. 2008) and a $t^{1/2}$ time-dependence based on variations in the $\Sigma_g - \dot{\Sigma}_\star$ relation with redshift (Tacconi et al. 2018). By imposing a floor at $\tau_\star = \tau_{\mathrm{mol}}$, our models do not allow the ISM to form stars more efficiently than molecular gas. We choose a threshold of $\Sigma_g = 10^8 M_\odot \mathrm{kpc}^{-2}$, because our model reaches these surface densities only in the inner ∼500 pc where the central molecular zone is found (Morris & Serabyn 1996; Dahmen et al. 1998; Pierce-Price et al. 2000; Hatchfield et al. 2020) and therefore where $\tau_\star = \tau_{\mathrm{mol}}$ would most plausibly occur.

3.1. *The Primary Set*

Potentially work references on MZR into this section. Motivate the choice of outflows over radial gas flows by calling out those papers. Our "primary set" of GCE models is so-named because they highlight the difference between equilibrium and "sub-equilibrium" GCE models. We use the term sub-equilibrium to emphasize the notion that the equilibrium metallicity can be mathematically defined for any choice of GCE parameters (see equation 5 below). The key question is not about the existence of an equilibrium state, but whether or not it is reached within the disk lifetime.

Following discussions of the MZR from the literature (e.g. Finlator & Davé 2008; Peeples & Shankar 2011), our primary set models differ first and foremost in their prescription for outflows from the Galactic disk. Weinberg et al. (2017) demonstrate that the equilibrium abundance of O in a one-zone GCE model is sensitive to the mass loading factor $\eta \equiv \dot{\Sigma}_{\mathrm{out}}/\dot{\Sigma}_\star$ relating the outflow and star formation rates according to

$$
Z_{\mathrm{O,eq}} = \frac{y_{\mathrm{O}}^{\mathrm{CC}}}{1 + \eta - r - \tau_\star/\tau_{\mathrm{sfh}}}.
\tag{5}
$$

The corrective term r accounts for the return of stellar envelopes back to the ISM, which can be approximated as a constant in most cases ($r \approx +0.4$ for a Kroupa 2001 IMF). In η is significant, then the factor of $\tau_\star/\tau_{\mathrm{sfh}}$ is a small correction, and the value of $Z_{\mathrm{O,eq}}$ is determined by stellar yields and the strength of outflows.

Equation 5 indicates that when $\eta \approx e^R$, the equilibrium abundance will decline roughly (though not exactly) exponentially with radius, tracking the observed metallicity gradient shape (e.g. Wyse & Silk 1989; Zaritsky 1992). We therefore adopt the following parameterization in two of our primary set models:

$$
\eta = \eta_\odot e^{(R-R_\odot)/R_\eta},
\tag{6}
$$

where $\eta_\odot$ sets the value of η at the solar radius ($R_\odot = 8$ kpc). The scale radius R_η of this relation approximately sets the slope of the equilibrium gradient $\nabla[\mathrm{O/H}]_{\mathrm{eq}}$ (see discussion in section 4.1 below) according to

$$
\nabla[\mathrm{O/H}]_{\mathrm{eq}} \approx \frac{-1}{R_\eta \ln 10}.
\tag{7}
$$
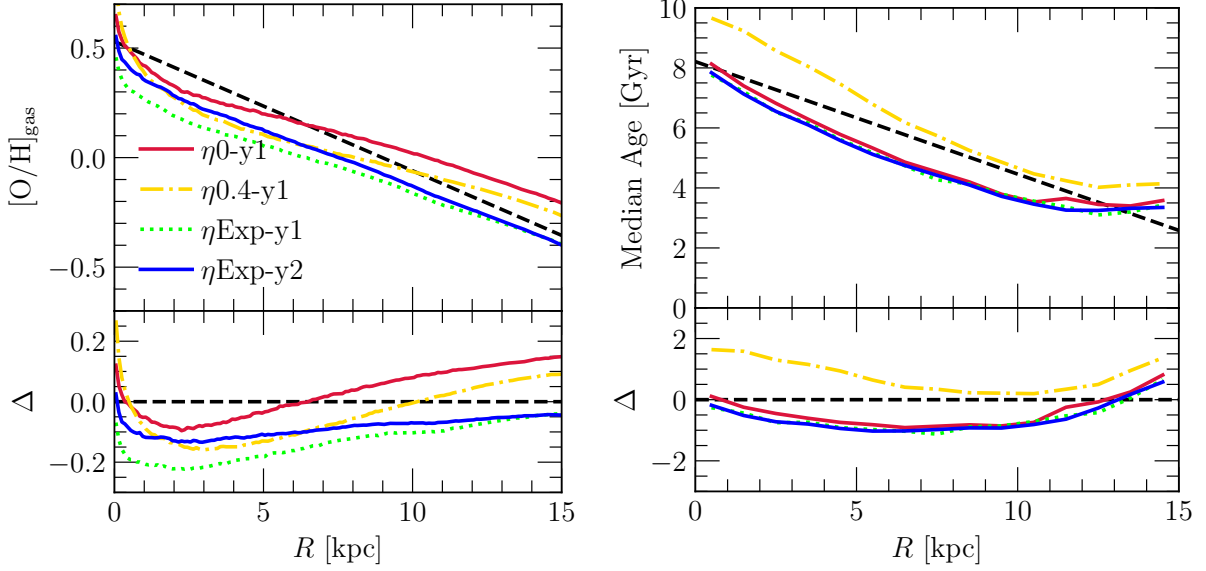
Figure 4. Calibration of our primary set of GCE models (see Table 2 and discussion in section 3 and Appendix B). **Left**: The predicted [O/H] gradients in the ISM at the present day (colored and styled lines marked by the legend). The black dashed line marks Méndez-Delgado et al.'s (2022) measurements in HII regions. The bottom panel shows the differences between the predicted and observed abundances as a function of radius. **Right**: The same as the left panel, but for the stellar age gradient. The black dashed line marks our linear regression in the bottom left panel of Fig. 2 (see fit parameters in Table 1). **Summary**: All models are calibrated to reproduce the observed ISM abundances at the present day, which they achieve sufficiently accurately for our purposes. Only the models invoking $\eta \propto e^R$ are calibrated to reproduce the observed age gradient, but those that are not still reasonably reproduce it.

We therefore use $R_\eta = 7$ kpc as a fiducial value, because this choice results in gradients consistent with our measurement of ∇[O/H] $= -0.062 \pm 0.001$ kpc^{-1} in section 2. Although η is highest in the outer disk, these models predict the surface density of the outflowing material to be highest at $R = 0$, because $\dot{\Sigma}_\star$ declines with radius faster than η increases.

This paper presents two models with the $\eta \propto e^R$ scaling, which we refer to as ηExp-$y1$ and ηExp-$y2$ based on their assumptions regarding η and the scale of stellar yields. The former assumes the $y/Z_\odot = 1$ normalization, while the latter uses the $y/Z_\odot = 3$ scale. These models also use a different normalization of the exponential scaling of η with radius,

namely $\eta_\odot = 0.4$ and $\eta_\odot = 2.4$, but both use the fiducial scale radius of $R_\eta = 7$ kpc. This additional parameter difference is necessary, because the overall scale of stellar yields is strongly degenerate with the strength of mass loading in GCE models (e.g. Cooke et al. 2022; Johnson et al. 2023c; Sandford et al. 2024).

We demonstrate in section 4 below that the ηExp-$y2$ model is a prototypical example of equilibrium chemical evolution, because it predicts ISM metallicities to reach a steady state nearly 10 Gyr ago. This prediction makes this model useful for highlighting the defining features of the equilibrium scenario. We therefore select ηExp-$y2$ as our fiducial set of

parameters. In section 3.3 below, we define a few simple variations of this base model.

As comparison cases, we also follow models with $\eta = 0$ and $\eta = 0.4$ at all radii (i.e. the limit that $R_\eta \to \infty$). Both of these models assume the $y/Z_\odot = 1$ scale of stellar yields. Following the same naming scheme, we refer to these models as $\eta0$-$y1$ and $\eta0.4$-$y1$. We highlight the $\eta = 0$ scenario as a particularly interesting comparison, because many GCE models of the MW assume no mass loading (e.g. Minchev et al. 2013, 2014; Spitoni et al. 2019, 2021; Palla et al. 2020, 2022; Gjergo et al. 2023). We discuss mass loading in GCE models of the MW and potential alternative origins of equilibrium chemical evolution in section 5.1 below.

3.2. *Parameter Calibration for the Primary Set*

Each model in our primary set closely approximates the present-day O abundance in the ISM by construction. In the $\eta0$-$y1$ and $\eta0.4$-$y1$ models, we tune the parameters τ_{rise} and τ_{sfh} describing the shape of the SFH (see equation 3) such that this observational result is reproduced. We describe this procedure in detail in Appendix B. In short, we first assume $\tau_{\mathrm{rise}} = 2$ Gyr as in J21 and search for a value of τ_{sfh} between 100 Myr and 200 Gyr that predicts the observed ISM abundance in an analytic one-zone GCE model. If no solution is found, we hold τ_{sfh} fixed at 200 Gyr and search for a value of τ_{rise} between 2 Gyr and $2\tau_{\mathrm{disk}} = 26.4$ Gyr. If still no solution is found, we simply use the values $\tau_{\mathrm{sfh}} = 200$ Gyr and $\tau_{\mathrm{rise}} = 26.4$ Gyr, which corresponds to an SFH that rises approximately linearly with time up to the present day.

In the ηExp-$y1$ and ηExp-$y2$ models, the present-day ISM abundances are much more sensitive to the exponential scaling of η with radius than the shape of the SFH (see discussion in section 3.1 above). These models therefore have more freedom than the $\eta = $ constant models in this regard, but at the expense of an additional free parameter. We therefore assign τ_{rise} and τ_{sfh} in the ηExp-$y1$ and ηExp-$y2$ models such that they reproduce the observed median stellar age as a function of Galactocentric radius. We describe this procedure in detail in Appendix B as well. A given combination of τ_{rise} and τ_{sfh} is accepted if the integral of the implied SFH up to the *observed* median age is equal to half of the integral up to the disk lifetime. Otherwise, we follow the same strategy as with the $\eta0$-$y1$ and $\eta0.4$-$y1$ models, starting with $\tau_{\mathrm{rise}} = 2$ Gyr and searching for a value of τ_{sfh} between 100 Myr and 200 Gyr. Although the observed age gradient is likely affected by selection effects in APOGEE, which we do not account for in this paper (see discussion in section 2.4), this uncertainty should not be a concern. The main goal is for the models to predict a plausible MW-like stellar disk but not necessarily a detailed match to the data.

Fig. 4 shows the results of this parameter calibration: the predicted ISM metallicities (left) and median stellar ages (right) as functions of Galactocentric radius for each model in our primary set. As intended, each model sufficiently reproduces the observed metal abundances in the ISM. Since the shape of the SFH is tuned to reproduce the observed stellar age in the ηExp-$y1$ and ηExp-$y2$ models, any deviations should be a consequence of radial migration. In our case, the

relation was altered from a purely linear trend to one that is slightly concave up. The observed age gradient is also reasonably approximated by the $\eta 0$-$y1$ and $\eta 0.4$-$y1$ models, even though they were not explicitly calibrated to agree with this trend.

In the η = constant models, our calibration procedure does not find a solution for τ_{rise} and τ_{sfh} at $R \gtrsim 7$ kpc and $R \gtrsim 10$ kpc, respectively. For these cases, we simply let the calibration procedure fail and adopt the values of $\tau_{\mathrm{sfh}} = 200$ Gyr and $\tau_{\mathrm{rise}} = 2$ Gyr. This failure arises because it is challenging to find parameters that keep the metallicity in the ISM as low as observed in the outer disk after 13.2 Gyr of star formation. As a result, these models overpredict the observed ISM abundance in these regions by a small amount, shown by the residuals in the bottom panels. In practice, these parameter choices are accurate enough for our purposes anyway, and calibration difficulties are not the reason for any of the successes or failures of these models. In our $\eta \propto e^R$ models, however, there are no radii at which the parameter calibration fails.

3.3. *Variations of the Fiducial Equilibrium Model*

Briefly call out slow and fast versions of the radial migration speed at the end of this section. Discussion of the ηevolExp-$y3$ model should obviously be removed if that model gets cut. I'm inclined to cut it, because it doesn't seem to get much discussion with the updates to the text, and the main point can be stated in words relatively concisely. Also state in words that we follow variations with 25% faster and 25% slower radial migration in section 5.2 and Appendix

C. Also call out the metal-rich accretion and metallicity dependent yield models in section 5.5.

Starting from the ηExp-$y2$ model, our fiducial choice of parameters, we construct four simple variations that lead to insightful comparisons. Two of these variations have different slopes, which we achieve by simply adjusting the value of R_η (see equation 6 and discussion in section 3.1 above). The ηExp-$y2$-steep model has a lower value of $R_\eta = 4.3$ kpc, which corresponds to a steeper slope of $\nabla[\mathrm{O/H}]_{\mathrm{eq}} = -0.1$ kpc^{-1}. The ηExp-$y2$-shallow is the opposite case, with a higher value $R_\eta = 10.9$ kpc and a shallower slope of $\nabla[\mathrm{O/H}]_{\mathrm{eq}} = -0.04$ kpc^{-1}.

Our third variations a model in which R_η suddenly changes at time $t = 8$ Gyr ($\tau = 5.2$ Gyr ago) from its initial value of 7 kpc to 21.7 kpc (i.e. $\nabla[\mathrm{O/H}]_{\mathrm{eq}} = -0.02$ kpc^{-1}). This ηevolExp-$y3$ model describes a scenario in which some property of the Galaxy changes in a manner that affects the long-term equilibrium state. We construct this model with no knowledge or assumption about what may cause such a change in nature. Our motivation is to predict what may happen if some intrinsic property of the Galaxy changes in a way that affects GCE parameters.

Our fourth and final comparison case is the ηExp-$y2$+burst model, which imposes a burst of star formation ~6 Gyr ago. The SFH in this model is given by

$$\dot{\Sigma}_\star \propto f_{\mathrm{rise\text{-}fall}}(t)\left[1 + A_b \Phi(t|t_b, \sigma_b, \alpha_b)\right], \qquad (8)$$

where Φ is a skew-normal function of time t centered on t_b with standard deviation σ_b and skewness α_b, and A_b is a dimensionless parameter describing the strength of the burst.
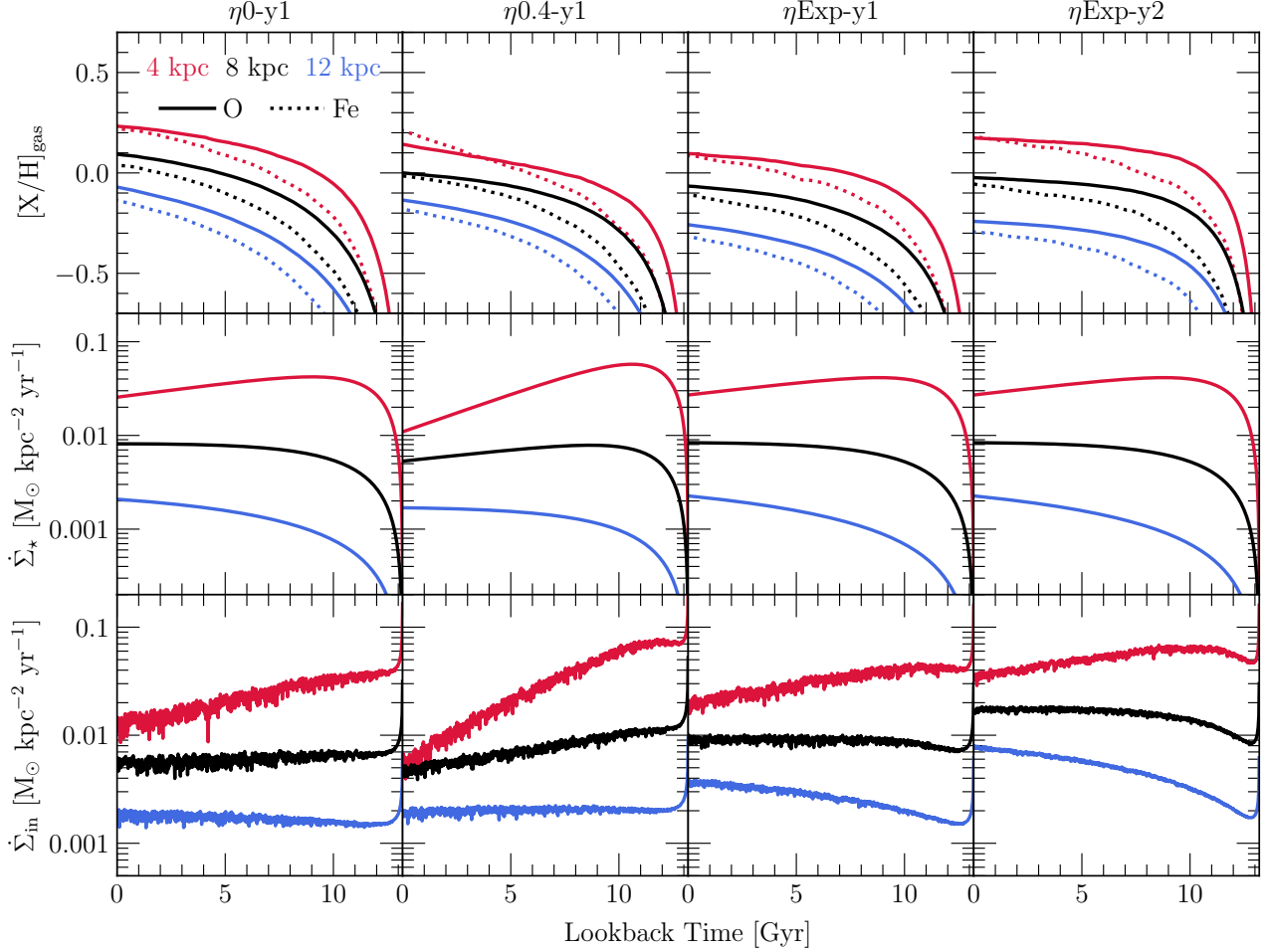
Figure 5. Enrichment (top), star formation (middle), and infall histories (bottom) for our primary set of GCE models, labeled at the top of each column of panels (see Table 2 and discussion in section 3). We visualize the evolution at Galactocentric radii of $R = 4$ (red), 8 (black), and 12 kpc (blue) in all panels. O and Fe abundances in the top row are distinguished with solid and dotted lines, respetively. **Summary**: Each model approximately reproduces the radial metallicity gradient observed in HII regions (Méndez-Delgado et al. 2022; see Fig. 4 in Appendix B), but evolves to that point differently due to differences in GCE parameters.

The model we present uses $t_b = 7$ Gyr, $\sigma_b = 1$ Gyr, and $\alpha_b = 3$, with A_b given by

$$A_b = \begin{cases} 0 & (R \leq 4 \text{ kpc}) \\ \min\left[2, e^{(R-4\text{ kpc})/5\text{ kpc}} - 1\right] & (R > 4 \text{ kpc}). \end{cases} \quad (9)$$

The burst starts small in amplitude at $R = 4$ kpc and grows in strength with radius, stopping at a factor of 3 increase in the SFR at $R \gtrsim 9$ kpc. We isolated these values by trial and

error, choosing them because they approximately reproduce the scale and timing of the variations in abundances seen in intermediate aged stars at $R \gtrsim 10$ kpc in Fig. 3.

4. RESULTS

Fig. 5 shows the enrichment, star formation, and accretion histories predicted by our primary set of models (see Table 2 and discussion in section 3.1). Each model reaches
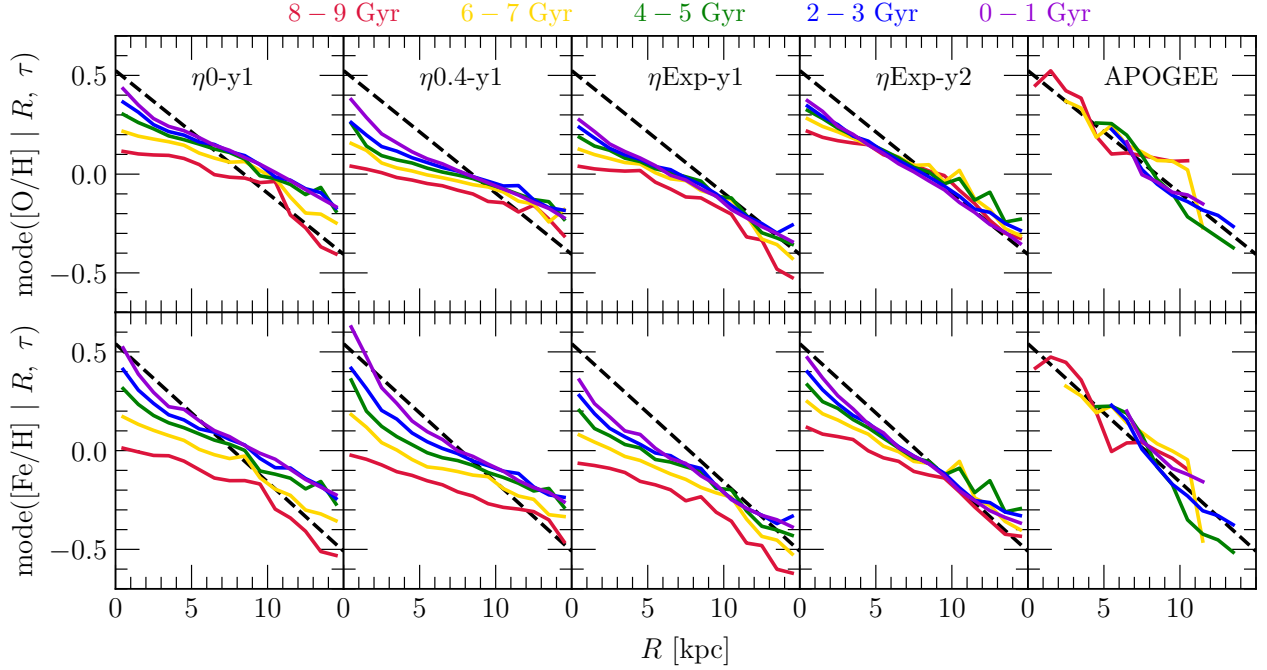
Figure 6. Comparing metallicity gradients in O (top) and Fe (bottom) for different ages of stellar populations (color coded according to the legend at the top left) between our measurements and our models. Stars mark our measurements with the ASTRONN value added catalog (see discussion in section 2) and are identical to the corresponding points in Fig. 3. Lines mark the corresponding measurements from our primary set of GCE models, denoted by the text in the top row of panels (see Table 2 and discussion in section 3). **Summary**: While agreement is not perfect, the ηExp-$y2$ model best reproduces the lack of evolution in the metallicity-radius relation with stellar age.

similar metal abundances in the ISM at the present day (see discussion in section 3.2 and Appendix B). What sets the models apart is that they reach these abundances on different timescales, which leaves observable signatures in the metallicities of mono-age populations.

The SFH in the outer disk is quite extended in the η0-$y1$ and η0.4-$y1$ models. This prediction arises due to our parameter calibration, which determines $\tau_{\rm rise}$ and $\tau_{\rm sfh}$ (see equation 3) such that the present-day ISM metallicities are closely approximated. A steadily increasing accretion rate keeps the metallicities low by ensuring a substantial supply of fresh hydrogen. Even though the ηExp-$y1$ and ηExp-$y2$ models have the same SFH, the accretion history is a factor of $\sim$3

higher in the latter. This difference is a direct consequence of the higher mass loading factor in the ηExp-$y2$ model. With stronger outflows from the ISM, more accretion is required to fuel a prescribed level of star formation and fulfill the stellar mass budget (see discussion in section 4.1.1 below).

To test these models against the observed metallicity-radius relation in mono-age populations, we apply the same procedure described in section 2.3 to our model predictions. The only difference is that we must weight the metallicity distributions in each bin of radius and age by the mass of each stellar population before computing the position of the mode. This additional step is necessary, because the mass of each
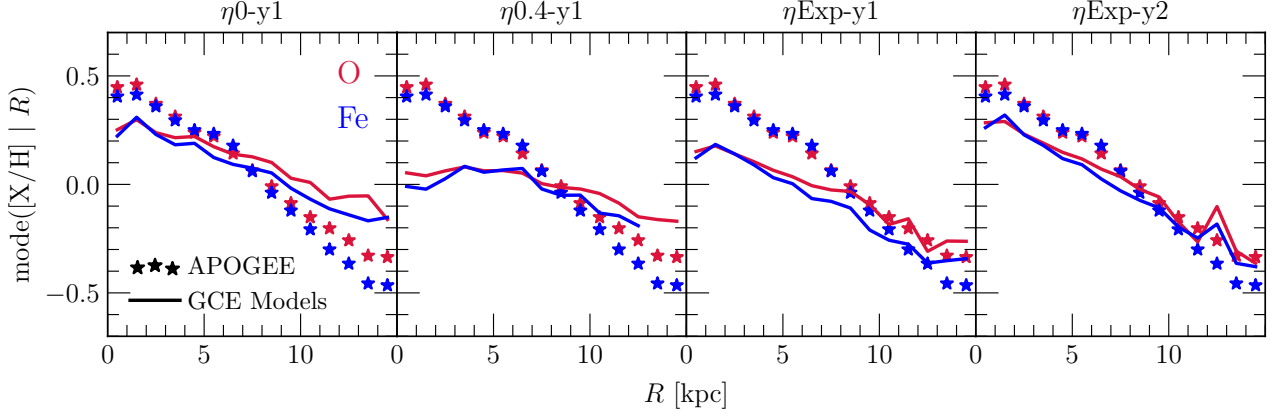
Figure 7. Comparing predicted stellar metallicity gradients with our empirical measurements. Stars denote our measurements from APOGEE in O (red) and Fe (blue) and are identical to the points in Fig. 2 (see discussion in section 2.3). Lines denote the predictions from our primary set of GCE models, marked at the top of each panel (see Table 2 and discussion in section 3). **Summary**: Since the models are calibrated to reproduce the observed gradient in the ISM, it is not guaranteed that they will also reproduce the observed stellar gradient. The ηExp-$y2$ model offers the best explanation of the consistency between gradients in gas and stars seen in Fig. 2.

stellar population determines the number of stars that would be available in an APOGEE-like survey.

Fig. 6 shows the results of this procedure for both O and Fe. The ηExp-$y2$ model best reproduces the empirical result that populations of different ages follow a similar relation between metallicity and Galactocentric radius. The agreement is not perfect; most notably, this model predicts some subtle differences in [Fe/H] between age bins at $R \lesssim 5$ kpc, but it performs sufficiently well for a qualitative reproduction. The η0.4-$y1$ and ηExp-$y1$ models perform reasonably well at $R \gtrsim 10$ kpc, but at smaller radii, there are significant variations between age bins that are not observed. The η0-$y1$ model performs the worst in this comparison. This model does not fail due to obvious discrepancies with the observed gradient slope, but rather because it predicts evolution in the normalization of the metallicity-radius relation, which is not observed.

In Fig. 7, we compare the predicted metallicity-radius relation for all stars (i.e. without sorting into age bins) with our measurements from the lower-right panel of Fig. 2. Because our models are calibrated to reproduce the present-day ISM abundances, it is not guaranteed that they will also reproduce the stellar gradient. This comparison therefore isolates the model that best reproduces the observed consistency between the metallicities of gas and stars as a function of radius (see discussion in section 2.2). Once again, the ηExp-$y2$ model performs the best. The η0-$y1$ model is a close second, with a slope that is only slightly shallower than that of ηExp-$y2$. The ηExp-$y1$ and η0.4-$y1$ models fail most noticeably in the inner disk, where they underpredict stellar abundances significantly more than the ηExp-$y2$ and η0-$y1$ models. Holistically, our ηExp-$y2$ model offers the best agreement with our data when Figs. 6 and 7 are both considered.

4.1. *The Equilibrium Gradient*

In this subsection, we discuss the physical origin of radially dependent equilibrium metallicities in the ISM by comparing between models. In section 4.1.1 below, we demonstrate that the decline in Z_{eq} with radius tracks a decline in the ratio of star formation per unit accretion $\dot{\Sigma}_\star/\dot{\Sigma}_{in}$ on long timescales. The metal production rate is most sensitive to $\dot{\Sigma}_\star$, while $\dot{\Sigma}_{in}$ determines the rate at which fresh H is added to the ISM.

Fig. 8 shows the predicted O abundance in the ISM as a function of radius in two recent snapshots and at the present day. We also plot the equilibrium abundance $[O/H]_{eq}$ as defined by equation 5 (using the values of $\tau_\star$ at the final snapshots). Our fiducial equilibrium model, ηExp-$y2$, shows little to no evolution across this time interval, because it approaches the steady state early in the disk lifetime (see discussion below). The other models in our primary set exhibit more significant evolution in the ISM abundances over this time interval. As expected, each model tends to predict faster evolution in the ISM metallicitiy when it is farther below the local equilibrium abundance.

The left and middle panels of Fig. 9 show the predicted O and Fe abundances in the ISM at $R = 8$ kpc as a function of lookback time. The models differ most importantly in the time dependence of the abundance evolution. The ηExp-$y2$ model predicts metal abundances to change only minimally over the last $\sim 8 - 10$ Gyr of disk evolution. Other models in our primary set predict the ISM metallicity to increase more gradually until the present day. This difference is more obvious in [O/H] than [Fe/H], because alpha elements reach the

local equilibrium abundance earlier than iron-peak elements due to the SN Ia DTD (Weinberg et al. 2017).

For comparison, we include the values of [O/H] and [Fe/H] determined by our linear regressions applied to mono-age populations in Fig. 3 (reported in Table 1). Our models predict the ISM metallicity at $R = 8$ kpc to be marginally (~ 0.1 dex) lower than observed in stars, with the exception of the $\eta 0$-$y1$ model $\lesssim 4$ Gyr ago. This systematic offset could be corrected by a small increase in our metal yields in all models, which would be comfortably within the range allowed by the uncertainties in stellar nucleosynthesis (see discussion in section 3). Each model provides a reasonable explanation for the metallicities observed in young stars. However, the data indicate that the age dependence is minimal at most up to populations as old as $\sim 8 - 10$ Gyr. This realization was also highlighted recently by Gallart et al. (2024). Although the agreement is not perfect, our ηExp-$y2$ model provides a better explanation than our other models for the seemingly age independent metallicities of stellar populations.

In the right panel of Fig. 9, we plot the "processing timescale[6]," defined as

$$\tau_{proc} \equiv \frac{\tau_\star}{1 + \eta - r}, \qquad (10)$$

as a function of radius at the present day. This quantity describes the average time interval that a given ISM fluid element will remain present before being incorporated into new stars or ejected in an outflow. In a one-zone model with

---

[6] In Johnson et al. (2021), we refer to this timescale as the "depletion time." We change nomenclature in this paper, because this term has other meanings in the literature.
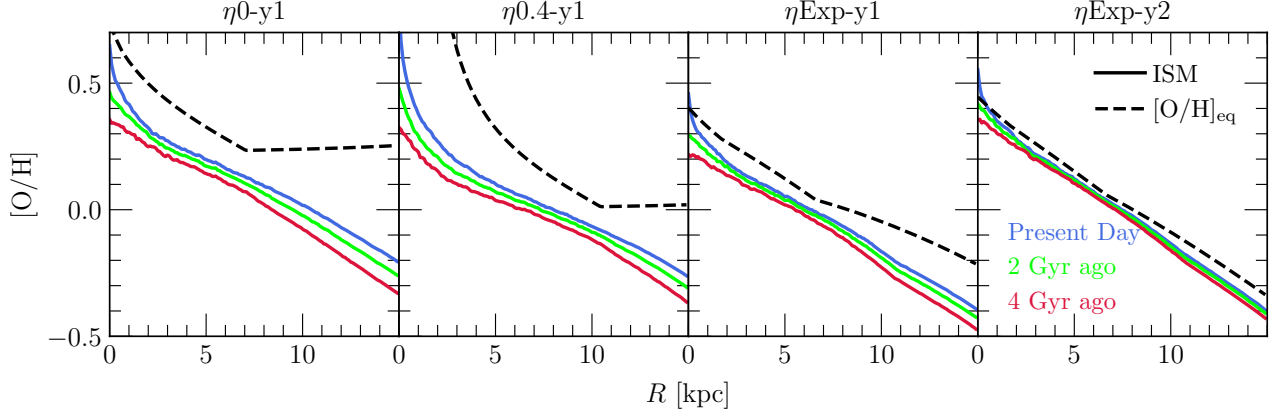
Figure 8. Evolution in gas-phase O abundance gradients in comparison to their equilibrium abundances. Solid lines show the abundances at the present day and at two recent snapshots predicted by each of the four GCE models in our primary set (see Table 2 and discussion in section 3). The black dashed line marks the equilibrium abundance as a function of radius R (see equation 5). **Summary**: The ηExp-$y2$ model is very near the equilibrium metallicity gradient, while other models do not reach this state within the disk lifetime.

a constant SFH, Weinberg et al. (2017) demonstrate that the ISM reaches the equilibrium alpha element abundance on this timescale:

$$Z_O(t) = Z_{O,eq} \left(1 - e^{-t/\tau_{proc}} \right). \qquad (11)$$

The corrective factor for most other forms of the SFH is negligible for interpretive purposes. The processing timescale can be qualitatively understood as setting the rate of approach to equilibrium, because the ISM loses some memory of its initial chemical composition every time it is effectively replaced with new matter. Consequently, an equilibrium only arises after a handful of "generations" of baryons in the ISM, the duration of which is set by τ_{proc}.

The important connection illustrated by Fig. 9 is that the ηExp-$y2$ model reaches high abundances more quickly than the other models, because it predicts the shortest processing timescales. There is a relationship between the value of τ_{proc} and the slope of [X/H] as a function of time in the ISM, with

shorter processing timescales corresponding to more constant abundances. The trend in τ_{proc} with radius is sometimes non-monotonic, because inefficient star formation slows down ISM processing at large radii, but strong outflows speed it up (see equation 10).

4.1.1. *The Ratio of Star Formation per Unit Accretion*

Fig. 10 shows the ratio of star formation per unit infall as a function of time at $R = 8$ kpc in our primary set of GCE models (see discussion in section 3.1). The maximum allowed value for a constant SFH is $1/(1 + \eta - r)$ (see equation 12 below), so each model asymptotically approaches a different value. The ηExp-$y2$ model reaches its maximum early in the disk lifetime, while other models follow suit on longer timescales.

There is information in both the time dependence of this ratio and its final value. In combination with the scale of stellar yields, the final value sets the local equilibrium abun-
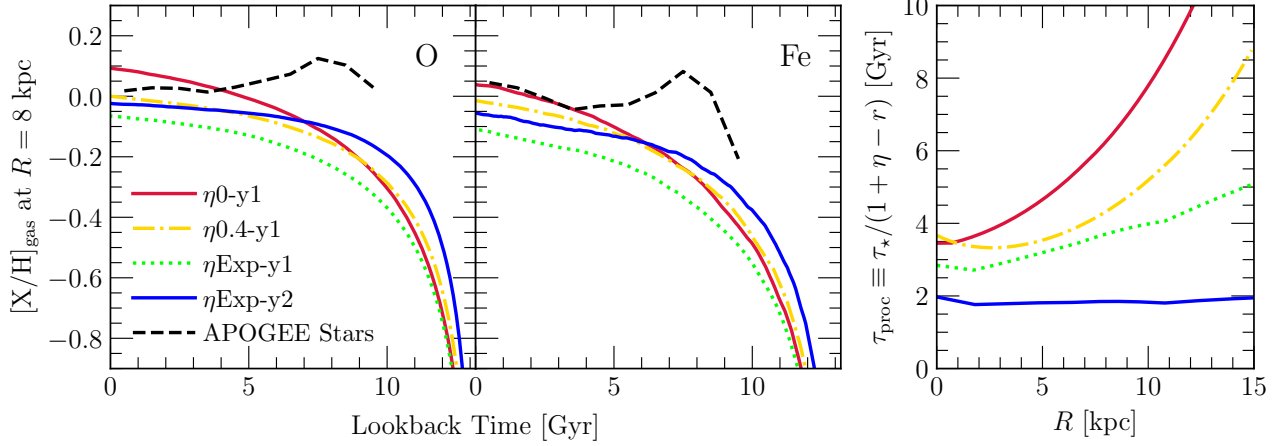
Figure 9. The key difference between our GCE models, marked according to the legend in the left panel. **Left/Middle**: The O (left) and Fe (middle) abundances in the $R = 8$ kpc ring as a function of time as predicted by our primary set of GCE models (see Table 2 and discussion in section 3). For comparison, we plot the values inferred from our linear regressions reported in Table 1, which are identical to the red (O) and blue (Fe) lines in the lower right panel of Fig. 3 (see discussion in section 3). **Right**: The processing timescale of the ISM (see equation 10) as a function of radius at the present day. **Summary**: Reaching high abundances early in the disk lifetime requires short processing timescales, meaning that a typical fluid element does not remain in the ISM for long ($\lesssim 1 - 2$ Gyr) before being incorporated into new stars or ejected in an outflow.

dance (see discussion below). The ηExp-$y2$ model predicts $\dot{\Sigma}_\star/\dot{\Sigma}_{\rm in}$ to be lower than the other models, but it reaches similar metallicities since the scale of stellar yields is also a factor of 2 higher. Once $\dot{\Sigma}_\star/\dot{\Sigma}_{\rm in}$ becomes time independent, the ISM reaches the local equilibrium metallicity. The ηExp-$y2$ model therefore reaches the equilibrium state so early in the disk lifetime, because this ratio is nearly constant for ~10 Gyr.

It is straightforward to demonstrate this connection mathematically. From equation B1 in Appendix B, it follows algebraically that $\dot{\Sigma}_\star/\dot{\Sigma}_{\rm in}$ can be expressed in terms of our GCE parameters as

$$\frac{\dot{\Sigma}_\star}{\dot{\Sigma}_{\rm in}} = \left[1 + \eta - r + \tau_\star \frac{\dot{\Sigma}_g}{\Sigma_g} \right]^{-1}, \qquad (12)$$

where we have also substituted in the SFE timescale $\tau_\star \equiv \Sigma_g/\dot{\Sigma}_\star$. The rate of change in the O abundance (see equation

B3) can then be expressed as

$$\dot{Z}_{\rm O} = \frac{y_{\rm O}^{\rm CC}}{\tau_\star} - \frac{Z_{\rm O}}{\tau_\star} \left(\frac{\dot{\Sigma}_{\rm in}}{\dot{\Sigma}_\star} \right). \qquad (13)$$

Furthermore, if $\tau_\star$ is constant and the late-time SFH declines exponentially, then $\dot{\Sigma}_g/\Sigma_g = -1/\tau_{\rm sfh}$, and the quantity in square brackets is equivalent to the denominator of equation 5 for $Z_{\rm O,eq}$. The equilibrium abundance can then be expressed as

$$Z_{\rm O,eq} = y_{\rm O}^{\rm CC} \frac{\dot{\Sigma}_\star}{\dot{\Sigma}_{\rm in}}. \qquad (14)$$

We emphasize that this expression is only accurate at $Z_{\rm ISM} \approx Z_{\rm eq}$. Otherwise, $\dot{\Sigma}_\star/\dot{\Sigma}_{\rm in}$ does not reflect the equilibrium state. Equation 5 is a more reliable expression for $Z_{\rm eq}$, because it is written in terms of the input GCE parameters.

The equilibrium gradient slope then follows by taking the logarithm of equation 14 and differentiating with respect to
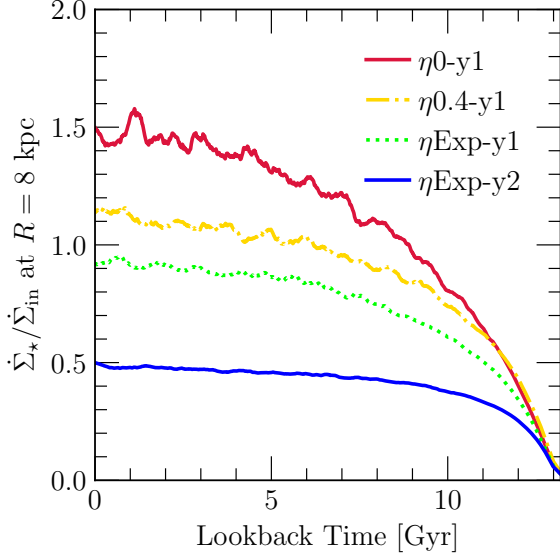
Figure 10. The star formation per unit infall at $R = 8$ kpc as a function of lookback time predicted by our primary set of GCE models. For visual clarity, we have box-car smoothed the trend with a window width of 200 Myr. **Summary**: The ISM can reach the local equilibrium abundance only after this ratio has reached some constant value (see discussion in section 4.1.1).

radius:

$$\nabla_{\text{eq}} = \frac{1}{\ln 10} \left[\frac{\partial \ln \dot{\Sigma}_\star}{\partial R} - \frac{\partial \ln \dot{\Sigma}_{\text{in}}}{\partial R} \right]. \qquad (15)$$

This expression neglects changes in the mass fraction of H, because it is typically a small correction (~0.01 dex). Due to the short lifetimes of massive stars (e.g. Larson 1974; Maeder & Meynet 1989; Henry et al. 2000), this relation is most accurate for alpha elements. The broad nature of the SN Ia DTD introduces an additional sensitivity to the shape of the SFH in the equilibrium abundance of Fe (Weinberg et al. 2017), which is likely related to why ∇[Fe/H] is slightly steeper than ∇[O/H] (see Fig. 2 and Table 1).

To demonstrate this relationship in action, we compare the ηExp-y2 model to the ηExp-y2-steep and ηExp-y2-shallow

variations (see discussion in section 3.3). Each of these models has the same value of $\eta_\odot$, but with different scale radii R_η describing how quickly the mass loading factor increases with radius. The left panel of Fig. 11 shows the present-day O abundance in the ISM as a function of radius in comparison to the equilibrium gradients predicted by these models (computed by evaluating equation 5 at all radii). Each model is near its equilibrium state at the present day by construction, but the equilibrium gradients have different slopes.

The right panel of Fig. 11 shows the predicted star formation and accretion rates as a function of radius at the present day. The SFR is the same between all three models, because it is specified as an input (see discussion in section 3). The inferred accretion rates differ as a consequence of changes in the mass loading factor. Relative to the fiducial model, an increase in the mass loading factor is accompanied by a larger accretion rate, because more mass is required to fuel star formation and fulfill the stellar mass budget of the Galactic disk when more ISM material is ejected. The ηExp-y2-shallow model at $R < 8$ kpc and the ηExp-y2-steep model at $R > 8$ kpc are examples of this prediction. Conversely, ηExp-y2-shallow at $R > 8$ kpc and ηExp-y2-steep at $R < 8$ kpc have lower accretion rates due to weaker mass loading than the fiducial model.

There is a clear inverse relation in the local accretion rate and metallicity at fixed SFR, which is a direct consequence of this connection. These ηExp-y2-shallow and ηExp-y2-steep models predict higher (lower) [O/H] than the fiducial model if the accretion rate is lower (higher). The effect of mass
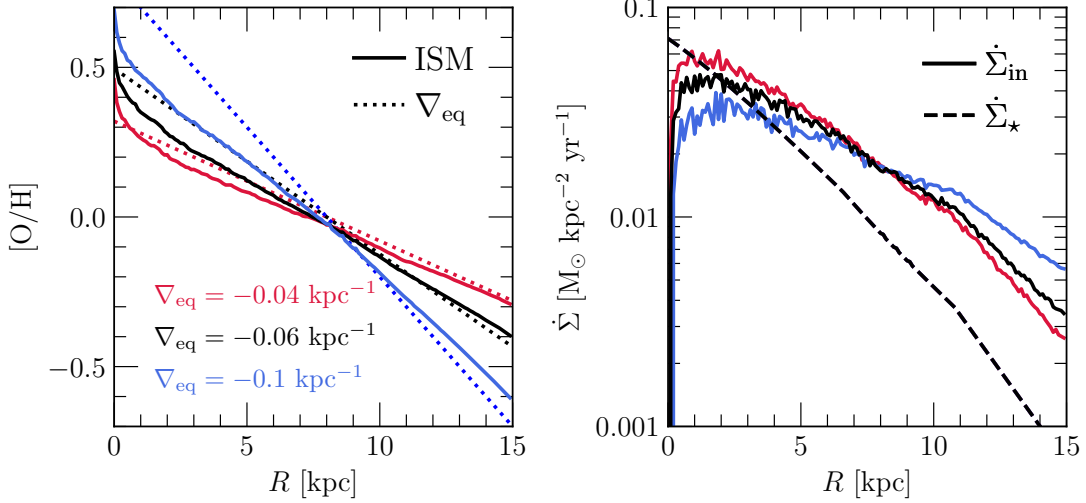
Figure 11. The relationship between the ratio of star formation per unit infall $\dot{\Sigma}_\star / \dot{\Sigma}_{in}$ and the slope of the equilibrium metallicity gradient ∇_{eq}. Models are color coded by the intended slope, which is controlled by the value of R_η (see Table 2), according to the legend in the right panel. **Left**: The ISM [O/H] abundance at the present day (solid) and the equilibrium gradients (dotted). **Right**: The surface densities of star formation (dashed) and infall (solid) at the present day as a function of Galactocentric radius. **Summary**: The slope of the equilibrium gradient is determined by variations in the time-averaged ratio of star formation per unit infall $\dot{\Sigma}_\star / \dot{\Sigma}_{in}$ with radius (see discussion in section 4.1.1).

loaded outflows on metal enrichment rates, as described by Weinberg et al. (2017) and supported by Fig. 11, is to remove metal-rich material from the ISM and replace it with metal-poor gas through accretion. This ongoing dilution lowers the equilibrium abundance. If this process happens in a radially dependent manner, then a radial gradient in the equilibrium metallicity arises. We note that we have found similar predictions when holding the accretion rate $\dot{\Sigma}_{in}$ fixed between models, in which case the surface density of star formation $\dot{\Sigma}_\star$ varies as a consequence of different choices in η.

4.2. *Perturbations and Sudden Events*

As additional explorations of the parameter space, we discuss the ηevolExp-$y3$ and ηExp-$y2$+burst models in this section (see discussion in section 3.3). The ηevolExp-$y3$ model is defined by a sudden change in R_η from a value of 7 kpc

to 21.7 kpc. As a result, the mass loading factor becomes suddenly larger at small radii and lower at large radii. The equilibrium gradient pivots around the solar abundance at $R = 8$ kpc to a new slope of $\nabla[O/H]_{eq} = -0.02$ kpc^{-2}.

Fig. 12 shows the initial and final scalings of η with radius alongside snapshots of the O abundance in the ISM immediately following the change in R_η at $t = 8$ Gyr. As expected, the ISM metallicity shifts gradually from its initial equilibrium state to the new equilibrium state on $\sim$Gyr timescales. The observational signature of such a model is a change in the gradient slope $\nabla[O/H]$ and $\nabla[Fe/H]$ in mono-age stellar populations. We discuss this prediction further in the context of recent arguemnts in the literature in section 5.3 below.

The ηExp-$y2$+burst model is defined by a burst in star formation in the central and outer disk (see equation 8) but
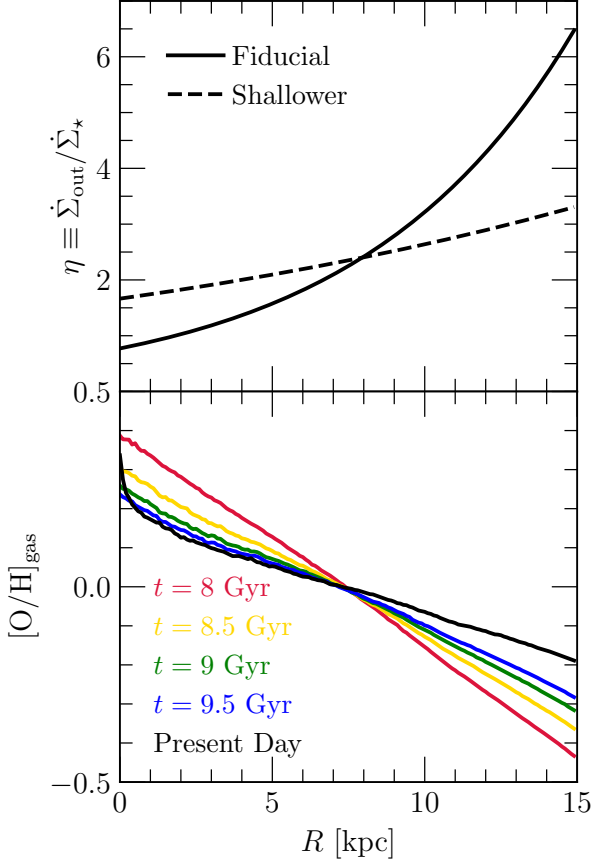
Figure 12. *It is probably worth cutting the ηevolExp-$y3$ model entirely as it does not seem to be discussed to significant depth with an updated version of the text. It can be stated in words somewhere that changing R_η results in a tilt of the gradient as expected.* The impact of building a sudden change in R_η at $t = 8$ Gyr into our ηExp-$y2$ model (see Table 2 and discussion in section 3). **Top**: The mass loading factor η as a function of radius before (solid) and after (dashed) the perturbation. **Bottom**: The [O/H] abundance in the ISM at the present day (black) and at a handful of snapshots immediately following the parameter change (lines color coded according to the legend). **Summary**: Parameter changes lead to changes in the equilibrium abundance with time, which can lead to shifts in the slope of the metallicity gradient. Though we do not find any obvious signs of such events in our sample, the literature has seen some recent arguments in favor of evolution in its slope (see discussion in section **??**).
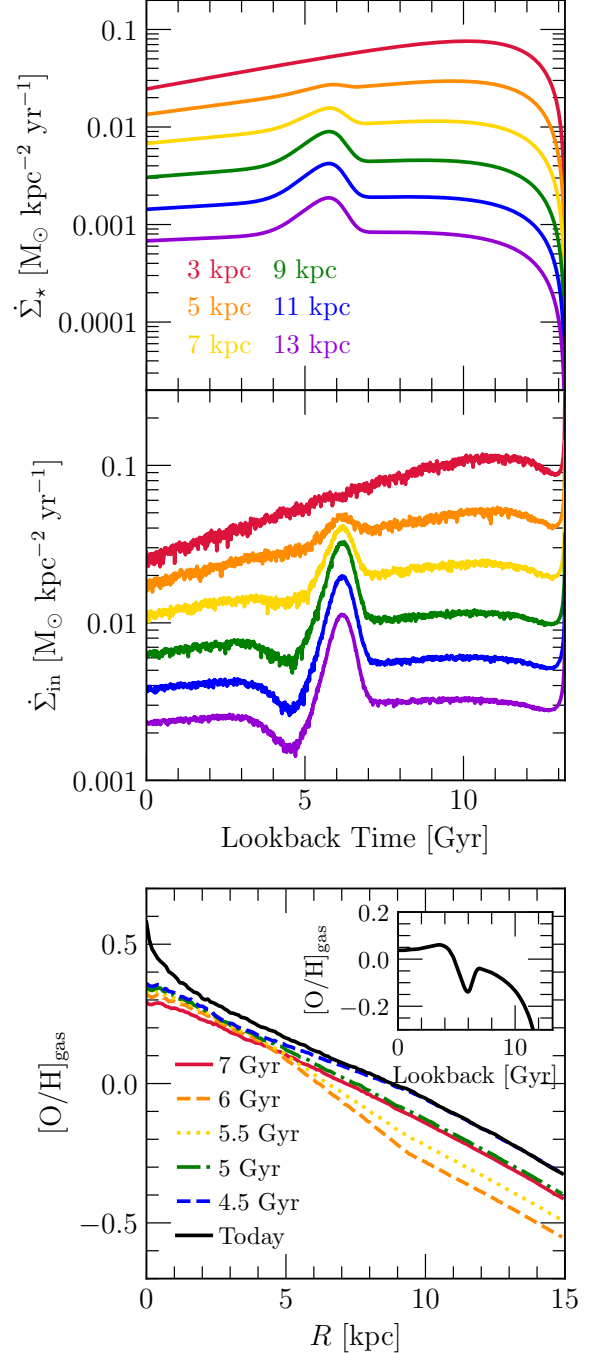
Figure 13. Our ηExp-$y2$+burst GCE model (see Table 2). **Top/Middle**: The input SFH (top) and predicted infall history (middle) at a selection of six Galactocentric radii, color coded according to the legend in the top panel. **Bottom**: [O/H] in the ISM as a function of Galactocentric radius, marked by the lookback time of the snapshot according to the legend. **Inset**: [O/H] in the ISM at $R = 8$ kpc as a function of lookback time. **Summary**: This model predicts the metallicity gradient to vary with a scale and timing that is broadly consistent with the $\sim 4 - 6$ Gyr populations in our sample (see Fig. 3 and discussion in section 2.3).

is otherwise the same as our fiducial model. The top and middle panels of Fig. 13 show the star formation and accretion histories at a selection of Galactocentric radii. In the outer disk, the accretion rates increase by a factor of a few for $\sim 1-2$ Gyr in order to fuel the enhanced star formation activity. The "jitter" in these predictions arises due to minor, stochastic fluctuations in the rate of return of stellar envelopes back to the ISM introduced by stellar migration, which affects the inferred level of accretion required to fuel the specified level of star formation (see discussion in section 4.1.1 above).

The bottom panel of Fig. 13 shows the O abundance in the ISM as a function of radius at a handful of snapshots covering the duration of the starburst. The ISM metallicity first decreases due to the dilution associated with the enhanced accretion rates. Following this time period, re-enrichment from star formation elevates the gas-phase abundances back to their pre-burst values within ~ 2 Gyr. The inset shows this prediction as a function of time at $R = 8$ kpc. The effects of this accretion induced burst of star formation can be understood as a perturbation that evolves back to the initial equilibrium state on $\sim$Gyr timescales, in line with previous theoretical arguments about dilution and re-enrichment (e.g. Dalcanton 2007; Johnson & Weinberg 2020). We also discuss this model further in the context of the literature in section X below.

5. DISCUSSION

In section 2.3, we demonstrated that stellar populations of different ages follow similar relations between metallicity and Galactocentric radius. While the slope of this relation is the focus of a lot of discussion in the literature, this paper is much more concerned with its normalization. In particular, old stars are not obviously more metal-poor than co-spatial young stars across much of the Galactic disk. This result is a recent realization that has been found independently with multiple methods for determining stellar ages (see discussion and references in section 1).

Of the GCE models that we focus on in this paper (see Table 2 and discussion in section 3.1), the one that reproduces this empirical result most readily is ηExp-$y2$. This model makes the following predictions, which we highlighted throughout section 4 as generic features that define the equilibrium model:

Feature 1. At all radii within the Galactic disk, the ISM metallicity evolves toward the local equilibrium abundance Z_{eq}. In the absence of major mergers or changes in GCE parameters, metal abundances do not evolve significantly with time once Z_{eq} has been reached, even with ongoing star formation.

Feature 2. The equilibrium abundance itself declines exponentially with radius. The equilibrium gradient ∇_{eq} refers to the slope of this relation, which is an intrinsic property of the Galactic disk set by variations in GCE parameters with radius.

Feature 3. The evolution toward equilibrium is fast ($\sim$Gyr timescales). Consequently, the equilibrium state is reached early in the disk lifetime.

Although we selected ηExp-$y2$ as a prototypical example of the equilibrium scenario, it is only one example. We

expect there to be alternative choices that would lead to similar predictions (see discussion in section X below).

Our ηExp-y2 model is quite similar to our fiducial model from J21. Therein we simply asserted the $\eta \propto e^R$ scenario, which fixed the radial metallicity gradient between models, allowing us to isolate the effects of different SFHs and radial migration timescales. We demonstrated therein that this class of models explains many details of the disk age-abundance structure. In particular, the dependence of the [O/Fe]-[Fe/H] distribution on Galactocentric radius and mid-plane distance (e.g. Hayden et al. 2015; Vincenzo et al. 2021) and the shapes of the age-[O/H], age-[Fe/H], and age-[O/Fe] relations (e.g. Feuillet et al. 2018, 2019) are readily explained. The most noteworthy shortcoming is that the latest versions of these models fail to reproduce the distinct bimodality seen in [O/Fe] distributions at fixed [Fe/H] (see also Dubay et al. 2024). We discuss this point further in section X below.

Paragraph: Discuss the burst and sudden change models. Main point of discussion is to emphasize that changes in the slope with time are allowed by the equilibrium scenario.

The local equilibrium abundance is reached quickly if baryons do not remain available for star formation for long ($\lesssim 1 - 2$ Gyr; see Fig. 9), because they are either incorporated into new stars or ejected in an outflow. As a consequence, the ratio of star formation per unit accretion $\dot{\Sigma}_\star/\dot{\Sigma}_{in}$ quickly reaches some constant value (see Fig. 10) describing the fraction of baryons that are ever incorporated into stars. The metal production rate $\dot{\Sigma}_Z$ is set by $\dot{\Sigma}_\star$, while $\dot{\Sigma}_{in}$ determines the rate $\dot{\Sigma}_H$ at which hydrogen is introduced into a star form-

ing region. The local equilibrium abundance Z_{eq} corresponds to the value at which these two rates balance, at which point the composition of the ISM becomes time-independent if the Galaxy is left undisturbed.

The equilibrium gradient arises out of a decline in $\dot{\Sigma}_\star/\dot{\Sigma}_{in}$ toward large radii (see Fig. 11). This ratio is sensitive to the mass loading factor η, which increases exponentially with radius in our fiducial model. In the outer disk, a smaller fraction of baryons form into stars, and by extension into metals, because a larger fraction are ejected in a wind. Based on this argument, we expect that the equilibrium scenario requires a process that removes material from the star forming ISM in a radially dependent manner. We continue this line of discussion in section 5.1 below.

Although our main empirical result is that stellar metallicities do not vary significantly across a surprisingly broad range of ages, we have not precisely measured how long ago the MW reached equilibrium. Our measurements in Fig. 3 indicate that $\sim 8 - 9$ Gyr old stars reflect the local equilibrium abundance across a broad range of radii, but the age uncertainties are substantial (see Fig. 1). Despite the age uncertainties, we believe that our empirical results are qualitatively sound. In order for the lack of evolution between bins in Fig. 3 to be driven entirely by difficulties in age estimation, the measurements need to be unreliable to the extent that they struggle to tell the difference between young stars and old stars. Stellar age measurements, although challenging indeed, are not uncertain to such an extent (see discussion in section 2.4).

5.1. *Mass Loaded Outflows Versus Radial Gas Flows*

Our GCE models demonstrate that the equilibrium scenario can arise due to wind-driven outflows with a mass-loading factor that increases exponentially with radius. Most importantly, the outflow affects the ratio of star formation per unit accretion $\dot{\Sigma}_\star/\dot{\Sigma}_{\rm in}$, which is closely related to the local equilibrium metallicity (see discussion in section 4.1.1). However, mass-loaded outflows are only one possible origin of the equilibrium scenario. Our models do not account for the effects of radial gas flows (e.g. Lacey & Fall 1985; Portinari & Chiosi 2000; Spitoni & Matteucci 2011; Grisoni et al. 2018). The flow of ISM material is thought to be directed inward (Bilitewski & Schönrich 2012; Vincenzo & Kobayashi 2020; Trapp et al. 2022). Such a prescription would continually replace local ISM material with lower metallicity gas from larger radii, qualitatively similar to the role of an outflow (Weinberg et al. 2017; see discussion in section 4.1.1). We therefore expect that radial gas flows may also influence the local equilibrium metallicity, potentially leading to a steady state evolution.

Although we do not differentiate between mass loaded outflows and radial gas flows in this paper, we can nonetheless make a handful of statements relatively confidently based on our results. First, we suspect that the equilibrium scenario requires at least one process that removes material from the star forming ISM. Our $\eta 0$-$y 1$ model, which does not eject any ISM material, does not predict the equilibrium gradient to be reached in the disk lifetime and is in tension with the observations (see Fig. 3). Second, the coefficient describing the rate of removal relative to star formation should increase

exponentially with radius. This criterion is a generalization of the $\eta \propto e^R$ prescription built into our fiducial model. With a different dependence on Galactocentric radius, the equilibrium gradient would have the wrong shape (see discussion in section 3.1). Lastly, this removal coefficient should not vary significantly on ~Gyr timescales for most of the disk lifetime. If we had let η depend on time in our models, then the equilibrium metallicity would not be constant, and we would expect variations in the ISM abundances, which violates the defining features of the equilibrium scenario (see discussion in section 5 above).

Based on recent results, it may be challenging for radial gas flows to pose a viable solution if a stable equilibrium state is required to explain the disk age-abundance structure. Doppler shifts of the 21 cm line indicate HI velocities spanning a broad range of values (Di Teodoro & Peek 2021). Some galaxies are dominated by inward gas flows, while some are dominated by outward gas flows, and others have both in different regions (see Appendix B of Di Teodoro & Peek 2021). Hydrodynamic simulations have been shown to predict a qualitatively similar diversity of gas velocities, but when averaged over time and azimuth, the speeds are of order a ~few km/s directed inward (Trapp et al. 2022). If the variability in radial gas velocities is substantial, it may be difficult for models to predict a stable equilibrium state in the absence of mass loaded outflows. Nonetheless, radial gas flows in the context of the equilibrium scenario is an interesting combination that we plan to investigate in future work.

There is also skepticism surrounding the viability of mass loaded outflows in GCE models of the MW disk, which plays a central role in our ηExp-$y2$ model (see discussion in section 3.1). Simulations of fountains predict that ejected metals should be re-accreted on short timescales ($\lesssim 100$ Myr; Spitoni et al. 2008, 2009) near the SN that produced them ($\lesssim 0.5$ kpc; Melioli et al. 2008, 2009). Many GCE models therefore adopt the limit of $\eta = 0$, assuming that the gravitational potential of the MW is too strong to launch a substantial outflow (e.g. Minchev et al. 2013, 2014; Spitoni et al. 2019, 2021; Palla et al. 2020, 2022; Gjergo et al. 2023). In some models, a portion of stellar yields is deposited directly into an outflow, but ambient ISM material is not ejected (e.g. Schönrich & Binney 2009; Chen et al. 2023).

While these arguments have motivated many GCE models to assume $\eta = 0$, there are other simulations in which mass loaded outflows do emerge from MW-like galaxies (e.g. Gutcke et al. 2017; Nelson et al. 2019; Kopenhafer et al. 2023). Hydrodynamic simulation predictions depend strongly on the adopted sub-grid models (e.g. Li et al. 2020) as well as the manner in which feedback and the grid itself are implemented (e.g. Hu et al. 2022). From a purely empirical perspective, outflows have never been observed from MW-like galaxies, because the predicted column densities are below the detection thresholds of current instruments (see discussion in e.g. the review by Veilleux et al. 2020). To date, evidence of mass loading has been observed only in strongly starbursting systems (e.g. M82, Lopez et al. 2020; NGC 253, Lopez et al. 2023; Mrk 1486, Cameron et al. 2021). Whether or not sig-

nificant outflows should emerge from more quiescent systems like the MW disk remains an open question.

All we can say with reasonable confidence is that GCE models omitting both mass loaded outflows and radial gas flows (e.g. our $\eta 0$-$y1$ models) fail to reproduce the age-independence of stellar metallicities across the Galactic disk (see Fig. 6). We have demonstrated that mass loaded outflows with $\eta \propto e^R$ are only one potential explanation thereof. Although we postulate the equilibrium scenario in this paper, we cannot make any strong claims about its detailed origin, because it remains to be seen if radial gas flows play an important role.

5.2. *Radial Migration*

Radial migration refers to a series of dynamical processes that impact the guiding center radii of stars, which can carry them several kpc away from their birth radius (e.g. Sellwood & Binney 2002). There is now a deep literature demonstrating that this process is ubiquitous in numerical simulations of disk galaxies (e.g. Roškar et al. 2008b,a; Loebman et al. 2011; Minchev et al. 2011; Bird et al. 2012, 2013; Kubryk et al. 2013). In this subsection, we discuss the role of radial migration in the equilibrium scenario.

The left panel of Fig. 14 shows distributions in [O/H] predicted by our ηExp-$y2$ model in three different ranges of Galactocentric radius. We include variations of this model in which radial migration is 25% slower ($\sigma_{RM8} = 2.0$ kpc; see discussion in section 3) and 25% faster ($\sigma_{RM8} = 3.6$ kpc; consistent with the value inferred by Frankel et al. 2018). We also determine the distribution that would arise without radial
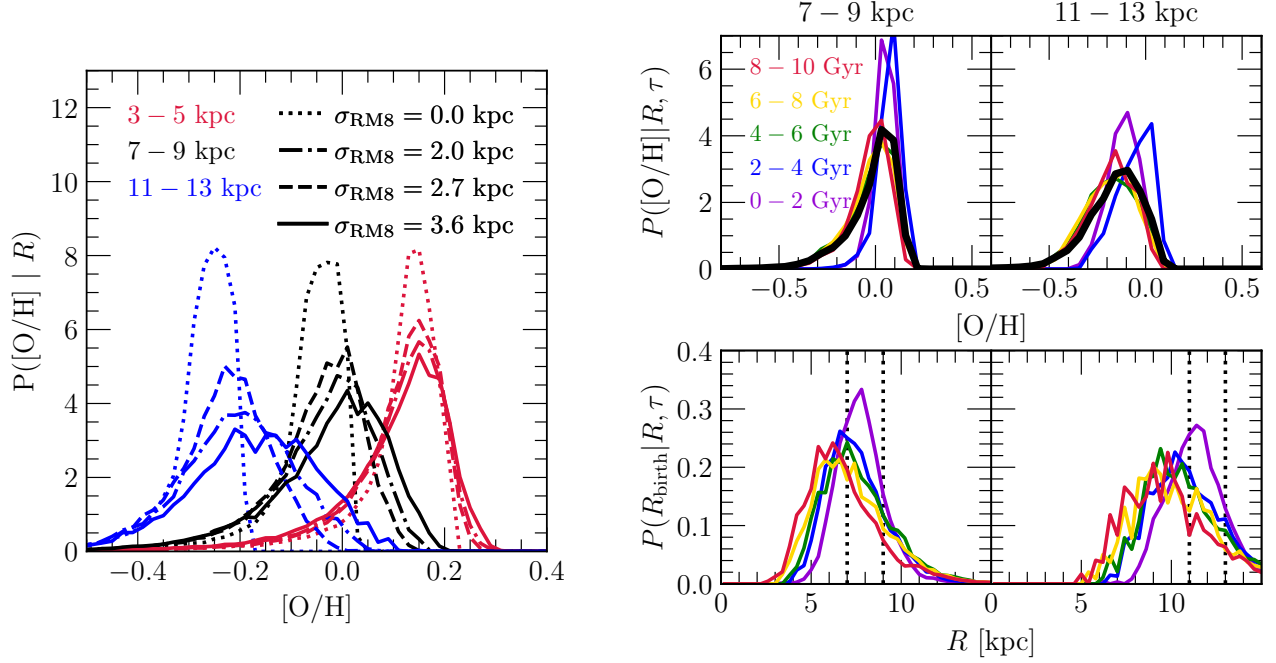
Figure 14. The impact of radial migration on abundance distributions. **Left**: Distribution in [O/H] in our ηExp-y2 model with different radial migration speeds σ_{RM8} (see discussion in section 3), styled according to the legend in the top right. Lines are color coded by their 2-kpc bin in present day radius according to the legend in the top left. **Right**: Distributions in [O/H] (top) and birth radius R_{birth} (bottom) at $R = 7 - 9$ kpc (left) and $R = 11 - 13$ kpc (right). Lines are color coded by stellar age according to the legend in the top left panel. The solid black line in the top panels shows the total [O/H] distribution. **Summary**: In the radial equilibrium model, stellar migration enhances the tails of the metallicity distribution but has only a minimal impact on the position of the peak, because the metallicity gradient is shallow compared to typical radial migration displacements (see equation 16).

migration ($\sigma_{RM8} = 0$) by simply sorting stars based on birth radius instead of final radius. In each model, mixing simply enhances the wings of the distribution but leaves the position of the peak largely unaffected. This prediction is in line with chemodynamic simulations by Grand et al. (2014).

It may seem counterintuitive that radial migration does not substantially affect the centroid of the MDF. At fixed present-day radius, there are many more stars that migrated outward than inward due to the steep nature of the surface density gradient (see, e.g., our dissection of the h277 star particles in

Fig. 1 of J21). Because of the radial metallicity gradient, one might expect a more significant shift to higher metallicities.

We address this point in the right hand panels of Fig. 14, which shows distributions in [O/H] and birth radius broken down by age in two of the radial bins from the left panel. The shift in $P(R_{birth}|R,\tau)$ toward smaller radii with increasing age is at the level we found in h277 in J21. The [O/H] distributions, however, are more age independent. The two distributions are approximately related by the local slope of the ISM metallicity gradient at the corresponding lookback

time according to

$$P(R_{\text{birth}}|R, \tau) \approx P([\text{O/H}]|R, \tau)\frac{\partial[\text{O/H}]}{\partial R_{\text{birth}}}, \qquad (16)$$

but only approximately so since all of these variables are correlated. Therefore, the reason stellar migration is predicted to only minimally affect the centroid of the MDF is because the radial metallicity gradient is shallow compared to typical migration distances. Based on our measurement of $\nabla[\text{O/H}] \approx -0.06 \text{ kpc}^{-1}$, a star that migrates $\Delta R = 3$ kpc outward is only ~0.2 dex more metal-rich than the stars that formed there.

Although the effect is slight, our GCE models do predict radial migration to flatten metallicity gradients over time. This prediction is seen most clearly in [O/H] in our ηExp-y2 model in Fig. 6. The ISM is near the equilibrium state when the stellar populations in that panel form, but the older populations follow a visibly flatter [O/H]-radius relation than the younger populations. This effect has also been seen observationally by Magrini et al. (2016), who compared O abundances in HII regions and planetary nebulae in M31, M33, M81, and NGC 300.

Herein lies our motivation for selecting the mode as the summary statistic in quantifying radial metallicity gradients in sections 2 and 4. The realization that stellar migration should not substantially affect the position of the peak of the MDF makes it a good proxy for the ISM abundance at a given radius and lookback time. We find this to be true in *all* of our GCE models, including η0-y1 and η0.4-y1. The mode of the abundance distribution for a mono-age population at fixed radius is typically no more than $\sim0.1 - 0.2$ dex above the ISM abundance at the corresponding time and place. This predic-

tion also holds if we increase the speed of radial migration (by adjusting the value of σ_{RM8}; see Frankel et al. 2018 or equation C1 of Dubay et al. 2024), though the mode becomes a slightly worse proxy.

5.3. *Birth Radius Inferences*

Recent work has sought to quantify the strength of radial migration in the Galactic disk by combining stellar ages and abundances with GCE models (e.g. Frankel et al. 2018; Minchev et al. 2018). However, the adopted GCE models generally predict the ISM abundances to rise continually with time. The success of the equilibrium scenario in section 4 therefore challenges the validity of the models these inferences are based on. Interestingly, the challenge is such that birth radius estimates should become *less* model dependent. If the ISM abundance at fixed radius does not evolve significantly with time, then metallicity is a proxy for birth radius on its own.

Inferring a star's birth radius directly from its metallicity under the assumption of the equilibrium scenario is a simple procedure. One simply inverts the radial gradient and maps metallicity back onto radius. This method does not provide particularly reliable birth radius estimates, because major mergers should briefly perturb the ISM metallicity from its equilibrium state (see Fig. 13 and discussion in section 4.2). Nonetheless, if the equilibrium model is accurate, then one should at least recover qualitative trends in radial displacement and age predicted by simulations (citation). We do indeed find expected results from this procedure. Namely, older populations have a broader distribution of radial dis-

placement ΔR. At fixed radius, old stars tend to have $\Delta R > 0$. We do not present these inferences here, because precise birth radius estimates are beyond the scope of this paper.

Recently, Lu et al. (2022) constructed a new and novel method for determining birth radii. Their method uses the range of metallicities in mono-age populations as a proxy for the metallicity gradient slope at a given lookback time. If the decline in ISM abundances toward large Galactocentric radii is monotonic, then wide (narrow) MDFs should arise from steep (shallow) gradients. They calibrated a relation between Range([Fe/H]) and ∇[Fe/H] based on hydrodynamic simulations, after which they infer birth radii for a sample of subgiants from LAMOST (Luo et al. 2015). Ratcliffe et al. (2023) subsequently used this method to infer birth radii for stars in APOGEE.

Both Lu et al. (2022) and Ratcliffe et al. (2023) argue that the metallicity gradient was steeper in the past based on the MDF broadening with age. At first glance, it may seem like this argument is in tension with our equilibrium chemical evolution scenario. However, their argument is centered on the slope of the relation between stellar metallicities and Galactic radius, while ours is centered on its normalization. The equilibrium scenario allows for changes in slope at fixed normalization. We have run additional variations of our ηExp-$y2$ model in which R_η, and by extension the equilibrium gradient, varies with time. The prediction is as expected; the metallicity gradient tilts over time, chasing after a moving equilibrium state.

We do however raise concerns about the reliability of birth radii inferred from this method. We discuss this point further in Appendix C. In short, our GCE models predict Range([Fe/H]) to increase by ~0.2 dex between 0 and ~10 Gyr even for a constant ∇[Fe/H]. This prediction is a direct consequence of radial migration. With increasing age, a given Galactic region samples a broader range of birth radii, and therefore a broader range of metallicity. However, this methodology for inferring birth radii assumes a one-to-one relationship between Range([Fe/H]) and ∇[Fe/H] as a function of age (see e.g. Fig. 2 of Lu et al. 2022 and/or Fig. 2 of Ratcliffe et al. 2023). A portion of the trend should instead be attributed to radial mixing itself. We suspect that a more precise understanding of the dynamical mixing history of the Galactic disk is required before birth radii can be reliably determined through the relationship between MDF widths and stellar age.

5.4. *Empirical Constraints on Stellar Yields*

If the equilibrium scenario is accurate, then trends in abundance ratios with metallicity should follow trends in stellar yields. This implication can be seen by inspection of equation 5. The abundance ratio of two elements A and B that have reached equilibrium can be expressed as

$$[A/B]_{eq} = \log_{10}\left(\frac{Z_{A,eq}}{Z_{B,eq}}\right) - \log_{10}\left(\frac{Z_{A,\odot}}{Z_{B,\odot}}\right) \tag{17a}$$

$$= \log_{10}\left(\frac{y_A(Z)}{y_B(Z)}\right) - \log_{10}\left(\frac{Z_{A,\odot}}{Z_{B,\odot}}\right), \tag{17b}$$

where we have generalized equation 5 to take into account metallicity dependent yields. Most importantly, the denom-

inator of equation 5, which depends on GCE parameters, cancels and leaves behind only the yield ratio.

In the equilibrium model is accurate, then the abundance structure of the thin disk therefore contains little information on its SFH. The result highlighted by Weinberg et al. (2019) that [X/Mg]-[Mg/H] trends do not vary with Galactic radius and mid-plane distance is expected. In exchange, these data contain a wealth of information on stellar yield ratios and should be regarded as viable testing grounds for models thereof. We found similar results in our previous investigation of N and O yields (Johnson et al. 2023b), while the detailed SFH is likely a source of scatter in abundance trends. In this paper, we are able to extend this argument to any element that has reached the local equilibrium abundance.

This implication of the equilibrium scenario is highly useful, because stellar nucleosynthesis is poorly understood. Nucleosynthesis is sensitive to many uncertain processes, including mass loss (e.g. Sukhbold et al. 2016), rotation and rotationally induced mixing (e.g. Frischknecht et al. 2016), and convection (e.g. Chieffi et al. 2001). Despite most massive stars living in binary systems (citation), stellar models have only recently started taking into account binary interactions (Farmer et al. 2021). Many massive stars may also collapse directly to black holes at the ends of their lives instead of exploding as CCSNe (e.g. Pejcha & Thompson 2015; Sukhbold et al. 2016; Ertl et al. 2016), which can have significant effects on population-averaged yields (Griffith et al. 2021). Even alpha and iron-peak elements, generally regarded as more well understood, require a "semi-empirical"
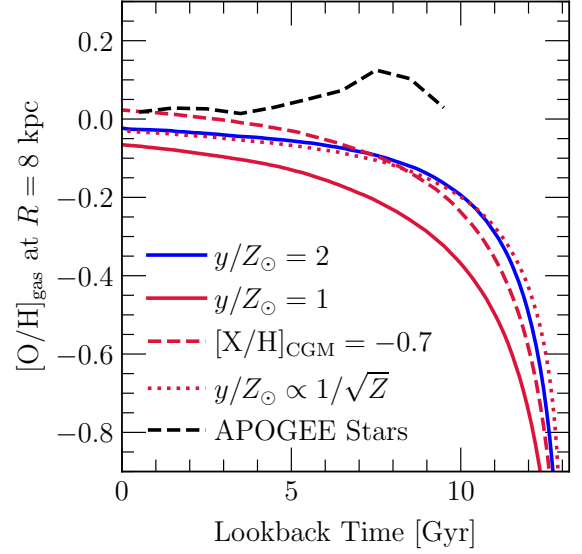
Figure 15. Variations in the ηExp-$y1$ model (all shown in red), which are intended to quicken its approach to equilibrium (see discussion in section 5.5 for details), with the ηExp-$y2$ model shown in blue. All curves are show the ISM [O/H] abundance as a function of lookback time in the $R = 8$ kpc annulus. The red dashed line shows a model in which the metallicity of accreting material is initially metal free but quickly reaches $[O/H]_{CGM} \approx [Fe/H]_{CGM} \approx -0.7$. The red dotted line shows a case in which we let all SN yields increase to low metallicity with a $1/\sqrt{Z}$ dependence based on recent arguments about SN event rates. The black dashed line is identical to the left panel of Fig. 9. **Summary**: Either stellar yields are considerably larger than recent arguments in the literature or the early SFH of the MW declined sharply with time (see discussion in section 5.5).

approach to accurately describe the evolution in the thin and thick disks (Palla et al. 2022).

5.5. *The Yield-Outflow Degeneracy*

The text of this section is based on the previous ηExp-$y3$ model with $y/Z_\odot = 3$. In that case, metal-rich accretion or metallicity dependent yields could *not* make up the difference

of a factor of 3 in stellar yields. However, the ηExp-$y2$ performs similarly well in the data-model comparisons in section 4. As it turns out, metal-rich accretion and metallicity dependent yields *can* make up a factor of 2 (see the updated version of Fig. 15). I therefore need to update the text in this section.

While observed abundance ratios should contain information on relative yields from stars (see discussion in section 5.4 above), constraining the absolute scale of stellar yields is much more challenging. GCE models make similar predictions when varying the yield normalization and the mass loading factor η in concert with one another (Johnson et al. 2023c; Sandford et al. 2024). However, our ηExp-$y1$ and ηExp-$y2$ models are sensitive to the difference, because the value of η influences the processing timescale τ_{proc}, which in turn influences how quickly the local equilibrium abundance is reached (see Fig. 9 and discussion in section 4.1). The result of Fig. 6 that our ηExp-$y2$ model best explains our data is therefore in tension with recent arguments from the literature. This model is on the scale of $y/Z_\odot = 3$ (see discussion in section 3). However, the observations by Rodríguez et al. (2022) based on the radioactive tails of CCSN light curves indicate a yield scale of $y/Z_\odot = 1$ (see discussion in Weinberg et al. 2023).

To investigate this discrepancy further, we construct two additional variations of our ηExp-$y1$ model. These variations are intended to "repair" the ηExp-$y1$ model so that it reaches higher metallicities at early times. In the first scenario, the metallicity of the CGM is initially zero but

follows a $1 - e^{-t/\tau}$ dependence and reaches a value of $[\mathrm{O/H}]_{\mathrm{CGM}} = [\mathrm{Fe/H}]_{\mathrm{CGM}} = -0.7$ on $\tau = 1$ Gyr timescales. In the second scenario, yields of all elements follow a $1/\sqrt{Z}$ dependence. Johnson et al. (2023a) argue that this metallicity dependence, which closely tracks the close binary fraction observed in APOGEE (Badenes et al. 2018; Moe et al. 2019), can explain the high SN Ia rates observed in dwarf galaxies (e.g. Brown et al. 2019; Wiseman et al. 2021). Pessi et al. (2023) recently showed that CCSNe also follow a strong inverse dependence on metallicity. Their inferred scaling is stronger than $1/\sqrt{Z}$, so this model is just a simple test to determine if higher SN rates overall can make up the difference between our ηExp-$y1$ model and the abundances observed in old stellar populations.

Fig. 15 compares the predicted [O/H] abundance at $R = 8$ kpc as a function of time between these variations, our ηExp-$y1$ and ηExp-$y2$ models, and the stellar abundances in our sample. Neither of these simple variations improve the ηExp-$y1$ model to the point that it agrees with the data. Both predict higher abundances at early times by construction, but not to the same extent as a factor of 3 overall increase in metal production.

To our knowledge, the only other way that high metal abundances can be reached sufficiently quickly is to make the early SFH decline sharply with time. This scenario can be understood as an initial epoch in which the equilibrium abundance is high because τ_{sfh} is low (see equation 5 and discussion in Weinberg et al. 2017). The more extended SFH of the thin disk then takes over, but the metallicity is already high, so

the new equilibrium state is reached quickly. This qualitative argument is indicative of an episodic SFH, such as the so-called "two-infall" model (e.g. Chiappini et al. 1997; Spitoni et al. 2019, 2021). We plan to investigate these evolutionary histories in the context of the equilibrium scenario in future work.

5.6. *Metallicity Gradients in the Cosmological Context*

Recent observations with JWST have indicated that the broad galaxy population is considerably more massive (Adams et al. 2023; Labbé et al. 2023; Naidu et al. 2022; Ferrara et al. 2023; Ferreira et al. 2022, 2023; Donnan et al. 2024) and more metal-rich than expected (Arellano-Córdova et al. 2022; Schaerer et al. 2022; Rhoads et al. 2023; Trump et al. 2023; Curti et al. 2023, 2024). These observations also support the empirical result that old stars in the MW are more metal-rich than expected (see Fig. 3 and discussion in section 2.4). Although we do not focus on high-redshift galaxies in this paper, the equilibrium scenario nonetheless offers a qualitative explanation when extended to other disks. Namely, we postulate that these galaxies are on their initial approach to chemical equilibrium. Our results indicate that ISM metallicities should evolve rapidly during this epoch.

The presence and ubiquity of radial metallicity gradients in the Universe is typically interpreted as evidence of inside-out disk growth (e.g. Matteucci & Francois 1989; White & Frenk 1991; Kauffmann 1996). A handful of GCE models have argued that a full explanation of the MW disk age-abundance structure requires an additional process acting in concert with inside-out growth (e.g. radial gas flows; Bilitewski

& Schönrich 2012; Grisoni et al. 2018; Palla et al. 2020). We support these arguments in this paper. In GCE models, inside-out growth is included via a more extended SFH (i.e. higher τ_{sfh} and/or τ_{rise}) and less efficient star formation (i.e. higher $\tau_\star$) at large radii. Our models indicate that at least one additional process is required in order to explain the high metallicities of old stellar populations. Although we use mass loaded outflows in this paper, we speculate that radial gas flows would make similar predictions (see discussion in section 5.1 above).

Our models instead indicate that the metallicity gradient reflects a decline in the time-averaged ratio of star formation per unit accretion toward large radii (see discussion in section 4.1.1). This decline is caused by the increase in the mass loading factor with radius, but future work is required to determine if radial gas flows have similar effects. Our $\eta \propto e^R$ scenario is *ad hoc*, because we chose this scaling knowing that it would predict a scaling of Z_{eq} with radius consistent with the observations (see discussion in section 3.1). It would nonetheless be reasonable to attribute this scaling to a relationship between η and $\Sigma_\star$. In other words, it becomes exponentially easier to launch a wind with increasing radius because the stellar surface density drops exponentially. If this argument is accurate, then metallicity gradients arise in spiral galaxies not because of inside-out disk growth, but because inside-out growth produces an exponential disk (e.g. Bird et al. 2013) whose gravitational potential leads to an $\eta \propto e^R$ profile.

An indirect relationship between inside-out disk growth and radial metallicity gradients lends a natural explanation

to their apparent universality. Disk galaxies exhibit a Gaussian distribution of slopes centered around ~-0.4 dex/R_{25} (~-0.1 dex/R_e) with a dispersion of about ~0.1 dex/R_{25} (e.g. Sanchez et al. 2014, Ho et al. 2015). There is no apparent correlation with galaxy morphology, including the presence of a bar (Sanchez-Menguiano et al. 2016). Some galaxies have radial gradients that invert in the innermost regions, or flatten in the outer regions, or both (Sanchez-Menguiano et al. 2018). Some authors have argued that the gradient slope is weakly correlated with galaxy stellar mass (Perez-Montero et al. 2014; Belfiore et al. 2017; Groves et al. 2023), while others have argued that there is no clear trend (Bresolin et al. 2019; Pilyugin et al. 2019; Poetrodjojo et al. 2018, 2021).

The variations in radial metallicity gradients between galaxies and across stellar mass is interestingly small. On the other hand, inside-out growth means something different to each individual spiral galaxy. Simulations predict a diversity of assembly histories (citation), with some experiencing mostly smooth and quiescent accretion while others undergo quite turbulent merger-driven growth. If there is a direct causal relationship between inside-out growth and metallicity gradients, then perhaps it should be surprising how little the radial trends vary between systems. A more natural explanation is that metallicity gradients are set by relatively universal galaxy properties. Equilibrium chemical evolution is a step in this direction.

6. CONCLUSIONS

Point to give emphasis to in revisions: the key differences between our GCE models lie in whether or not the age gradient and the metallicity gradient are driven by the same parameters or not. The best performing model, ηExp-$y2$, has the metallicity gradient set by *different* parameters than those that set the age gradient, which suggests they are set by different processes. Should also cite the Sharda et al. (2021) paper (also one from 2024?) here. Their model, which is much more first principles, also predicts an equilibrium state to arise. I think this conclusion is at a quite reasonable length, but there is probably room for one or two more paragraphs if necessary.

In this paper, we present the equilibrium model of disk galaxy metallicity gradients. Referring to a class of models rather than one specific choice of parameters, the defining feature of this scenario is that the observed ISM metallicity as a function of radius tracks a decline in the equilibrium metallicity Z_{eq}. Abundances asymptotically approach Z_{eq} and become nearly time independent after the first few Gyr of disk evolution.

The primary line of evidence for this scenario is that old stellar populations are more metal-rich than predicted by conventional GCE models (Willett et al. 2023; Gallart et al. 2024; see Fig. 6 and discussion in section 4). Recent observations with JWST indicate that high-redshift galaxies are also more metal-rich than expected (citation). If our Galaxy is a typical spiral, then this result supports the empirically motivated argument that the young MW should have been metal-rich (see discussion in section 2.4). Equilibrium GCE addresses these newfound discrepancies qualitatively.

We measure metallicity gradients of $\nabla[\text{O/H}] = -0.062 \pm 0.001 \text{ kpc}^{-1}$ and $\nabla[\text{Fe/H}] = -0.070 \pm 0.003 \text{ kpc}^{-1}$ for all stars

in our sample (see Fig. 2 and Table 1), in broad agreement with previous inferences (e.g. Chen et al. 2003; Daflon & Cunha 2004; Cheng et al. 2012; Frinchaboy et al. 2013; Myers et al. 2022; Spina et al. 2022; da Silva et al. 2023). When splitting our sample into 1-Gyr bins of age, these mono-age populations indicate that the slope settled at ∇[O/H] $\approx$ −0.05 kpc^{-1} and ∇[Fe/H] $\approx$ −0.06 kpc^{-1} ~9 Gyr ago (see Fig. 3). Stars in the ~4 − 6 Gyr age range follow steeper gradients of ∇[O/H] $\approx$ −0.07 kpc^{-1} and ∇[Fe/H] $\approx$ −0.09 kpc^{-1}. Young populations follow shallower gradients of ∇[O/H] $\approx$ −0.05 kpc^{-1} and ∇[Fe/H] $\approx$ −0.07 kpc−1, resembling the older populations.

The observation that most clearly favors the equilibrium scenario is that these variations in the *slope* of the relation between stellar metallicity and Galactocentric radius occur at nearly constant *normalization*. The differences in this relation traced by mono-age populations are at a level accurately described by an equilibrium state that was perturbed by a merger event ~6 Gyr ago (see Fig. 13 and discussion in section 4.2). Such an event can likely be attributed to accretion of metal-poor gas due to the first pericentric passage of the Sagittarius dwarf spheroidal (e.g. Law & Majewski 2010; Ruiz-Lara et al. 2020). GCE Models that predict the ISM metallicity to grow with time also predict changes in the normalization of this metallicity-radius relation between mono-age populations, in tension with the data (see Fig. 6 and discussion at the beginning of sections 4 and 5).

The success of the equilibrium scenario in reproducing the age-abundance structure of the MW disk has a number of implications for the broader picture of galactic astrophysics, which we discuss throughout section 5. At the heart of these implications is the notion that abundances and abundance ratios of thin disk stars contain much less information on the Galactic SFH than previously thought. Instead, equilibrium GCE predicts the data to much more closely trace trends in stellar yields with metallicity (see equation 17b). This prediction opens up a potentially highly useful avenue for empirically testing elemental yields predicted by stellar evolution models, which are poorly understood (see discussion in section 5.4).

We argue that inside-out disk growth (e.g. Matteucci & Francois 1989; White & Frenk 1991; Kauffmann 1996) is indirectly rather than directly related to radial metallicity gradients. In our GCE models, variations in the shape of the SFH and the efficiency of star formation with radius do not sufficiently explain the disk age-abundance structure. We require at least one additional parameter in order to explain the data. Our fiducial equilibrium model (ηExp-y2; see discussion in section 3.1) is successful with mass loaded outflows. We plan to explore radial gas flows (e.g. Lacey & Fall 1985; Portinari & Chiosi 2000; Spitoni & Matteucci 2011; Bilitewski & Schönrich 2012) in future work, with particular attention to whether or not they are also a possible origin of the equilibrium scenario.

APPENDIX

A. THE LEUNG ET AL. (2023) AGE CATALOG

In this appendix, we replicate our results in Fig. 3 with the Leung et al. (2023) catalog of stellar ages. We have focused our discussion in this paper on the AstroNN catalog (Mackereth et al. 2019), but it is possible that the neural network simply learned the correlations between chemical abundances and ages. In this case, our key empirical result in Fig. 3 would be subject to considerable systematic uncertainties. Leung et al. (2023) demonstrated that their estimates do not contain any significant amount
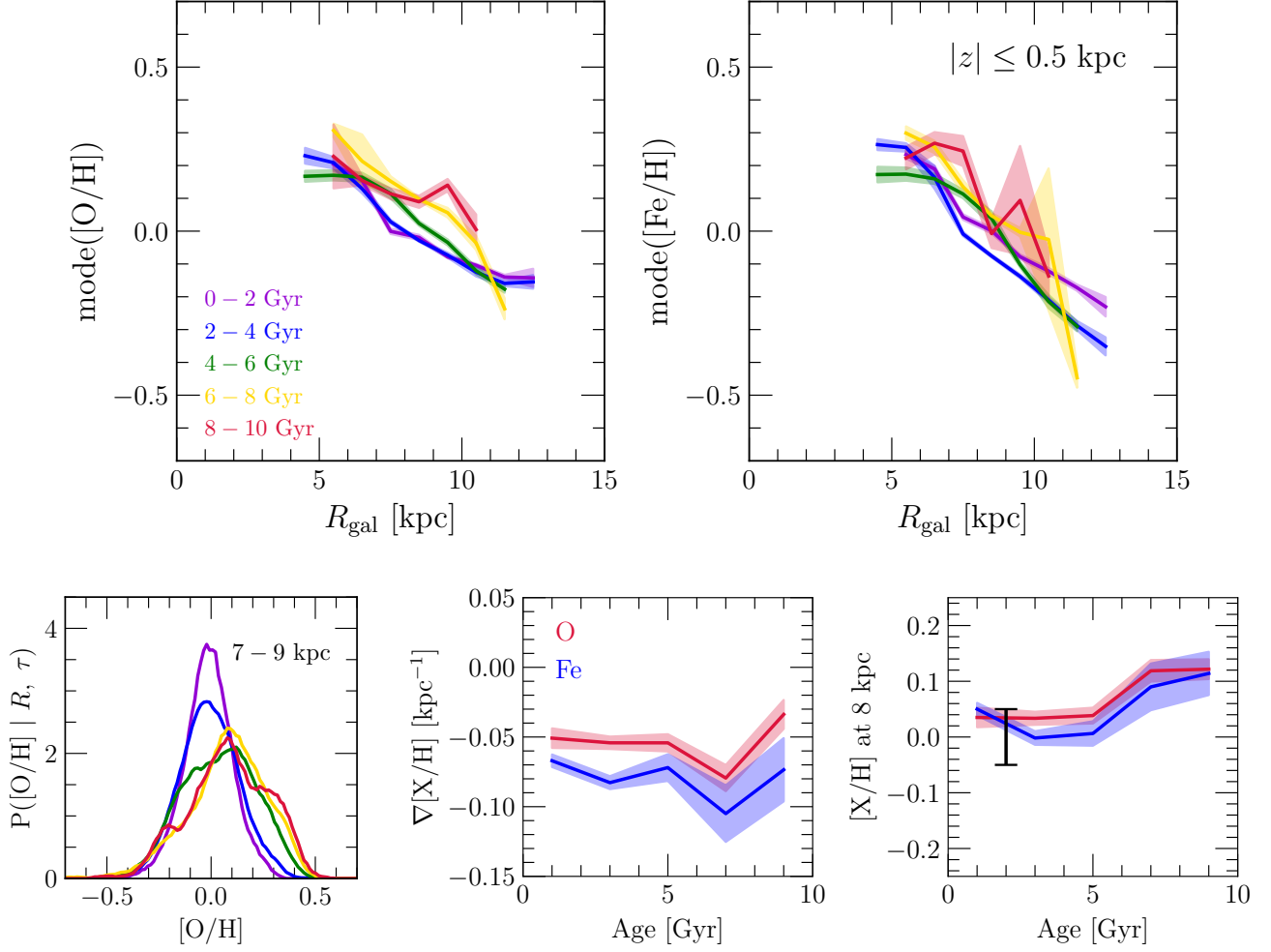
Figure 16. The same as Fig. 3, but using Leung et al.'s (2023) catalog of stellar ages as opposed to the AstroNN value added catalog. Because of the smaller sample, we use 2 Gyr bins in age as opposed to 1 Gyr according to the legend in the top left panel. **Summary**: We find the same results as in Fig. 3, so our central results should not be affected by any learned correlations between stellar abundances and ages in the AstroNN catalog.

of information on stellar abundances (see discussion in section 2.4), making their catalog an excellent comparison case to validate our measurements.

They mitigate this potential issue by compressing the spectra into lower dimensional representations of themselves (i.e., a *latent space*) using a variation encoder-decoder algorithm (e.g., LeCun et al. 2015). They then train a modified random forest algorithm to predict similarly compressed lightcurves trained on *Kepler* photometry Borucki et al. (2010). They demonstrate that this latent space contains little if any information on chemical abundances and stellar parameters, as intended, after which they are able to estimate ages by augmenting the latent space with $T_{\rm eff}$ and $\log g$ and decompressing the lightcurves.

Table 4. A summary of linear regression parameters applied to the [O/H]-R and [Fe/H]-R relations for mono-age populations in the Leung et al. (2023) catalog (see discussion in Appendix A).

Age Range	∇[O/H]	[O/H](R = 8 kpc)	∇[Fe/H]	[Fe/H](R = 8 kpc)
$0 - 2$ Gyr	-0.051 ± 0.007 kpc^{-1}	0.035 ± 0.017	-0.067 ± 0.004 kpc^{-1}	0.050 ± 0.011
$2 - 4$ Gyr	-0.054 ± 0.004 kpc^{-1}	0.034 ± 0.012	-0.083 ± 0.005 kpc^{-1}	-0.002 ± 0.012
$4 - 6$ Gyr	-0.054 ± 0.006 kpc^{-1}	0.038 ± 0.014	-0.072 ± 0.010 kpc^{-1}	0.006 ± 0.022
$6 - 8$ Gyr	-0.079 ± 0.009 kpc^{-1}	0.119 ± 0.019	-0.105 ± 0.020 kpc^{-1}	0.090 ± 0.042
$8 - 10$ Gyr	-0.034 ± 0.010 kpc^{-1}	0.122 ± 0.018	-0.073 ± 0.023 kpc^{-1}	0.114 ± 0.039
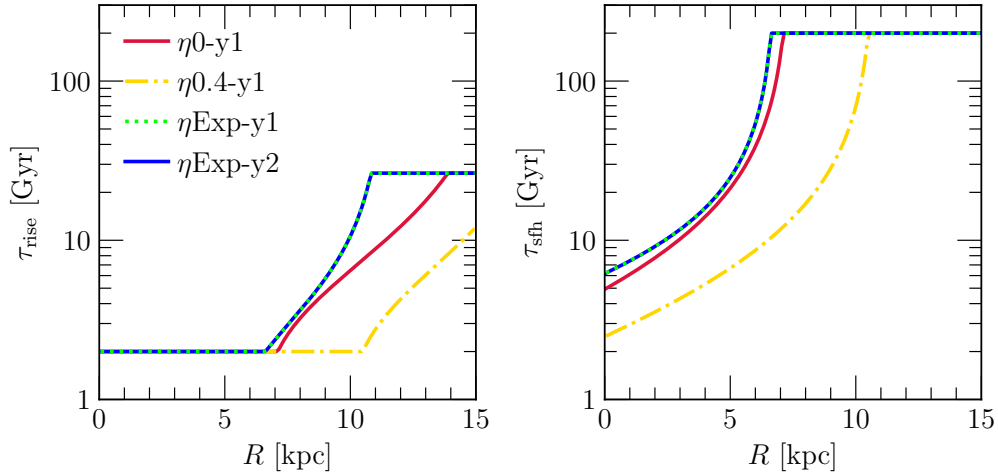
Figure 17. Our SFH parameter calibration. Curves show τ_{rise} (left) and τ_{sfh} (right) as a function of radius in each model, marked according to the legend in the left panel.

The Leung et al. (2023) training sample spanned a much narrower range in surface gravity ($\log_{10} g = 2.5 - 3.6$), which lowers the sample size by a factor of ~2.5. We therefore use 2 Gyr as opposed to 1 Gyr bins in age and compute the mode for subsamples containing at least 100 as opposed to 200 stars. Fig. 16 shows the results, with the lines of best fit to the metallicity gradients in each age bin reported in Table 4. As in Fig. 3, we find that both the slope and normalization of the disk metallicity gradient are independent of age up to ~10 Gyr, the full range of ages that we probe here. We can then safely conclude that our central results are not affected by any learned correlations between chemical abundances and stellar ages in the AstroNN catalog.

B. GCE MODEL PARAMETER CALIBRATIONS

In this appendix, we describe our procedure for assigning τ_{rise} and τ_{sfh} as functions of Galactocentric radius in our GCE models. These parameters describe the shape of the SFH in a given ring (see equation 3). Our procedure follows the analytic one-zone GCE models by Weinberg et al. (2017).

In a given annulus within the MW disk, the time derivative of the gas surface density[7] follows a summation of source and sink terms:

$$\dot{\Sigma}_g = \dot{\Sigma}_{\text{in}} - \dot{\Sigma}_{\star} - \dot{\Sigma}_{\text{out}} + \dot{\Sigma}_r$$
$$\approx \dot{\Sigma}_{\text{in}} - \dot{\Sigma}_{\star} \left(1 + \eta - r\right), \tag{B1}$$

where $\dot{\Sigma}_r \approx r\dot{\Sigma}_{\star}$ is the rate of return of stellar envelopes back to the ISM (see discussion in section 3.1). Our multi-zone models also follow equation B1 under the caveat that stellar populations may be exchanged between rings, so the exact rate of return varies a small amount with time. In fact, this stochastic exchange of mass is the reason for the "jitter" in the predicted accretion rates seen in Figs. 5, 10, and 11.

The rate of change in the surface density of O is given by equation 7 of Weinberg et al. (2017):

$$\dot{\Sigma}_O = y_O^{\text{CC}}\dot{\Sigma}_{\star} - Z_O\dot{\Sigma}_{\star}\left(1 + \eta - r\right). \tag{B2}$$

There is an additional factor of Z_O in the second term, because losses of O to star formation and outflows occur at the ISM abundance at a given time. The source term $y_O^{\text{CC}}\dot{\Sigma}_{\star}$ describes CCSN production of O under the approximation of instantaneous production, which is accurate enough for our purposes due to the short lifetimes of massive stars (e.g. Larson 1974; Maeder & Meynet 1989; Henry et al. 2000). The rate of change of the O abundance $\dot{Z}_O$ then follows from combining equations B1 and B2 with quotient rule:

$$\dot{Z}_O = \frac{\Sigma_g\dot{\Sigma}_O - \Sigma_O\dot{\Sigma}_g}{\Sigma_g^2}$$
$$= \frac{y_O^{\text{CC}}}{\tau_{\star}} - \frac{Z_O}{\tau_{\star}}\left(1 + \eta - r + \tau_{\star}\frac{\dot{\Sigma}_g}{\Sigma_g}\right), \tag{B3}$$

where we have also substitued in the SFE timescale $\tau_{\star} \equiv \Sigma_g/\dot{\Sigma}_{\star}$.

We hold $\tau_{\star}$ constant in time for the purposes of this parameter calibration. In our numerical models, we use the full time-evolution described by equation 4. Under this assumption, the rates of change in the SFR and the gas supply are related by $\ddot{\Sigma}_{\star}/\dot{\Sigma}_{\star} = \dot{\Sigma}_g/\Sigma_g$. Differentiating our rise-fall SFH (see equation 3) with time and plugging in the result yields the following

---

[7] The area of each ring in our multi-zone models are constant, so surface

density follows the same continuity equations as mass.

expression for $\dot{Z}_O$:

$$\dot{Z}_O = \frac{y_O^{CC}}{\tau_\star} - \frac{Z_O}{\tau_\star}\left(1 + \eta - r - \frac{\tau_\star e^{-t/\tau_{rise}}}{\tau_{rise}\left(1 - e^{-t/\tau_{rise}}\right)} - \frac{\tau_\star}{\tau_{sfh}}\right). \tag{B4}$$

This expression is a linear ordinary differential equation, whose solution has a known form (see equation 33 of Weinberg et al. 2017), which in this case is given by

$$Z_O(t) = Z_{O,eq}\left[1 - \exp\left(-t\frac{\tau_{sfh} - \tau_{proc}}{\tau_{sfh}\tau_{proc}}\right) - \frac{\tau_{sfh}\tau_{rise}}{\tau_{sfh}\tau_{rise} - \tau_{proc}\tau_{rise} - \tau_{sfh}\tau_{proc}}\left(e^{-t/\tau_{rise}} - \exp\left(-t\frac{\tau_{sfh} - \tau_{proc}}{\tau_{sfh}\tau_{proc}}\right)\right)\right], \tag{B5}$$

where $Z_{O,eq}$ is the equilibrium abundance (see equation 5), and τ_{proc} is the processing timescale (see equation 10). The integration constant is assigned such that $Z_O = 0$ at $t = 0$.

In our η0-y1 and η0.4-y1 models, we assign each ring values of τ_{rise} and τ_{sfh} such that equation B5 predicts the O abundance observed in the ISM at the present day. Following Fig. 2, we use the measurements by Méndez-Delgado et al. (2022) as our empirical benchmark. As described in section 3.2, we first assume $\tau_{rise} = 2$ Gyr and search for values of τ_{sfh} between 0.1 and 200 Gyr satisfying these criteria. If no solution is found, we adopt $\tau_{sfh} = 200$ Gyr and search for values of τ_{rise} between 2 Gyr and $2\tau_{disk} = 26.4$ Gyr. If still no solution is found, we simply adopt $\tau_{rise} = 26.4$ Gyr, which corresponds to an SFH that rises approximately linearly until the present day. For the sake of this parameter calibration, we use our scaling of $\tau_\star$ with gas surface density at the present day (see equation 4) assuming a simple exponential disk with a scale radius of $R_g = 3.75$ kpc (Kalberla & Kerp 2009).

In our ηExp-y1 and ηExp-y3 models, the present-day ISM metallicities are much more sensitive to the choice of mass loading factor η than the shape of the SFH. For these models, we therefore choose τ_{rise} and τ_{sfh} such that the 50th percentile of the integrated SFH matches the observed median age at a given radius:

$$\int_0^{\tau_{1/2}(R)}(1 - e^{-t/\tau_{rise}})e^{-t/\tau_{sfh}}dt = \frac{1}{2}\int_0^{\tau_{disk}}(1 - e^{-t/\tau_{rise}})e^{-t/\tau_{sfh}}dt. \tag{B6}$$

We otherwise follow the same procedure as described above, starting with $\tau_{rise} = 2$ Gyr and searching for a value of τ_{sfh}. We find a value of τ_{sfh} satisfying this criterion at all radii in these models, so there is no need to adopt $\tau_{sfh} = 200$ Gyr and search for a value of τ_{rise}.

Fig. 17 shows the resulting values of τ_{rise} and τ_{sfh} as a function of radius in our primary set models. The η0-y1 and η0.4-y1 models are forced to adopt the "fail-safe" values described above at $R \gtrsim 7$ kpc and $R \gtrsim 10$ kpc, respectively. Nonetheless, these models reasonably reproduce the observed radial gradients in both metallicity and age anyway (see Fig. 2). The issues with calibration in these instances arise because there is a limited region of parameter space that allows ISM metallicities to be as low as the observations in the outer disk. However, these difficulties are the not reason for the shortcomings of these models described throughout section 4.
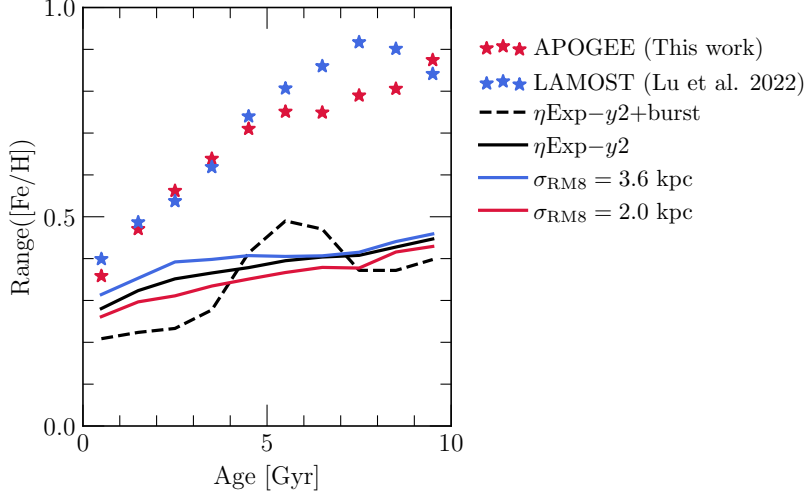
Figure 18. A comparison between our models and the methodology presented by Lu et al. (2022) for estimating birth radii for stars (see discussion in section 5.3). Range([Fe/H]) refers to the difference between the 5th and 95th percentiles of the MDF, shown here in 1-Gyr bins of age for stars between $R = 7$ and 10 kpc. Stars mark Lu et al.'s (2022) measurements using data from LAMOST (blue) and the same measurements for our sample from APOGEE (red). Lines, color coded and styled according to the legend, mark the predictions by our ηExp-$y2$ model and variations thereof (see discussion in sections 3.1 and 3.3). **Summary**: The relationship between Range([Fe/H]) and is influenced to some extent by stellar migration. Modeling these effects would improve the Lu et al. (2022) methodology.

C. BIRTH RADIUS INFERENCES

In section 5.3, we discussed the recent method put forth by Lu et al. (2022) for determining Galactocentric birth radii based on the relationship between the width of the [Fe/H] distribution and stellar population age. Therein we raised concerns about the reliability of this methodology. We discuss this point further in this appendix.

To compare the arguments by Lu et al. (2022) with our models, we apply their methodology to our predictions. To mimic measurement uncertainty, we perturb the ages and metallicities of each stellar population by random numbers drawn from normal distributions of width 0.5 Gyr and 0.03 dex, respectively. We then isolate stellar populations between $R = 7$ and 10 kpc at the present day and take Range([Fe/H]) to be the difference between the 5th and 94th percentiles of the [Fe/H] distribution for a mono-age population. We apply the same procedure to our sample from APOGEE but omit the perturbation by artificial uncertainties.

Fig. 18 shows the results of this procedure as a function of stellar population age in comparison to the LAMOST subgiants from Lu et al. (2022). The APOGEE stars follow a similar trend as LAMOST. The two surveys are in good agreement with one another, considering systematic uncertainties that can arise when comparing optical and infrared surveys with different radial coverage of the disk. Ratcliffe et al. (2023) also used APOGEE stars, but with ages measured with XGBoost (Miglio et al. 2021), a supervised

machine learning algorithm based on spectroscopically inferred stellar parameters and asteroseismic ages. Our measurements are in reasonable qualitative agreement with theirs, but with some differences in detail (see their Fig. 2).

We show Range([Fe/H]) as a function of age predicted by our ηExp-$y2$ and ηExp-$y2$+burst models along with the variations that use 25% faster and slower radial migration from section 5.2. Our models predict systematically narrower MDFs than observed. This discrepancy does not affect our main conclusions, because the arguments in this paper are based on the position of the mode of the distribution as opposed to its detailed shape. One potential origin of this discrepancy is short timescale (of order ~100 Myr) variations in the ISM metallicity, which should in principle arise due to the clumpy and episodic nature of accretion and star formation. Such effects are not present in our GCE models in this paper, because we use smooth SFHs.

Despite a nearly constant slope ∇[Fe/H], the ηExp-$y2$ model and its variations each predict Range([Fe/H]) to increase by ~0.2 dex between young and ~10 Gyr old populations. This prediction is a consequence of radial migration. The $R = 7 - 10$ kpc range, which this comparison focuses on, samples a broader range of birth radii with increasing age. Due to the presence of a radial abundance gradient, this region therefore also samples a broader range of metallicities. The trend depends minimally on the strength of radial mixing, but there is a trend nonetheless. As mentioned in section 5.3, we therefore raise concerns regarding the reliability of birth radii determined with this method. A portion of this trend should be attributed to radial migration, and therefore variations in Range([Fe/H]) and ∇[Fe/H] with age should have different shapes.

We nonetheless support the argument by Lu et al. (2022) that Range([Fe/H]) should increase and ∇[Fe/H] should decrease briefly during a merger event. Our ηExp-$y2$+burst model predicts Range([Fe/H]) to increase by ~$0.1 - 0.2$ dex at a time coincident with the enhanced accretion associated with this model. It is therefore reasonable to infer that the GSE merger event (citation) would have simultaneously steepened the Galactic abundance gradient and increased MDF widths ~$8 - 12$ Gyr ago. These effects are not apparent in our measurements with APOGEE, but that discrepancy could be an issue of systematic differences between the surveys.

REFERENCES

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35, doi: 10.3847/1538-4365/ac4414

Adams, N. J., Conselice, C. J., Ferreira, L., et al. 2023, MNRAS, 518, 4755, doi: 10.1093/mnras/stac3347

Aller, L. H. 1942, ApJ, 95, 52, doi: 10.1086/144372

Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140, doi: 10.1088/0004-637X/765/2/140

Andrievsky, S. M., Luck, R. E., Martin, P., & Lépine, J. R. D. 2004, A&A, 413, 159, doi: 10.1051/0004-6361:20031528

Arellano-Córdova, K. Z., Berg, D. A., Chisholm, J., et al. 2022, ApJL, 940, L23, doi: 10.3847/2041-8213/ac9ab2

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481, doi: 10.1146/annurev.astro.46.060407.145222

Badenes, C., Mazzola, C., Thompson, T. A., et al. 2018, ApJ, 854, 147, doi: 10.3847/1538-4357/aaa765

Beaton, R. L., Oelkers, R. J., Hayes, C. R., et al. 2021, AJ, 162, 302, doi: 10.3847/1538-3881/ac260c

Bedding, T. R., Huber, D., Stello, D., et al. 2010, ApJL, 713, L176, doi: 10.1088/2041-8205/713/2/L176

Bedding, T. R., Mosser, B., Huber, D., et al. 2011, Nature, 471, 608, doi: 10.1038/nature09935

Bigiel, F., Leroy, A., Walter, F., et al. 2010, AJ, 140, 1194,
    doi: 10.1088/0004-6256/140/5/1194

Bilitewski, T., & Schönrich, R. 2012, MNRAS, 426, 2266,
    doi: 10.1111/j.1365-2966.2012.21827.x

Bird, J. C., Kazantzidis, S., & Weinberg, D. H. 2012, MNRAS,
    420, 913, doi: 10.1111/j.1365-2966.2011.19728.x

Bird, J. C., Kazantzidis, S., Weinberg, D. H., et al. 2013, ApJ, 773,
    43, doi: 10.1088/0004-637X/773/1/43

Blanc, G. A., Lu, Y., Benson, A., Katsianis, A., & Barraza, M.
    2019, ApJ, 877, 6, doi: 10.3847/1538-4357/ab16ec

Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529,
    doi: 10.1146/annurev-astro-081915-023441

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977,
    doi: 10.1126/science.1185402

Bovy, J. 2016, ApJ, 817, 49, doi: 10.3847/0004-637X/817/1/49

Brown, J. S., Stanek, K. Z., Holoien, T. W. S., et al. 2019, MNRAS,
    484, 3785, doi: 10.1093/mnras/stz258

Cameron, A. J., Fisher, D. B., McPherson, D., et al. 2021, ApJL,
    918, L16, doi: 10.3847/2041-8213/ac18ca

Casamiquela, L., Tarricq, Y., Soubiran, C., et al. 2020, A&A, 635,
    A8, doi: 10.1051/0004-6361/201936978

Chaplin, W. J., & Miglio, A. 2013, ARA&A, 51, 353,
    doi: 10.1146/annurev-astro-082812-140938

Chartab, N., Newman, A. B., Rudie, G. C., Blanc, G. A., & Kelson,
    D. D. 2023, arXiv e-prints, arXiv:2310.12200,
    doi: 10.48550/arXiv.2310.12200

Chen, B., Hayden, M. R., Sharma, S., et al. 2023, MNRAS, 523,
    3791, doi: 10.1093/mnras/stad1568

Chen, L., Hou, J. L., & Wang, J. J. 2003, AJ, 125, 1397,
    doi: 10.1086/367911

Cheng, J. Y., Rockosi, C. M., Morrison, H. L., et al. 2012, ApJ,
    746, 149, doi: 10.1088/0004-637X/746/2/149

Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765,
    doi: 10.1086/303726

Chieffi, A., Domínguez, I., Limongi, M., & Straniero, O. 2001,
    ApJ, 554, 1159, doi: 10.1086/321387

Chieffi, A., & Limongi, M. 2013, ApJ, 764, 21,
    doi: 10.1088/0004-637X/764/1/21

Cooke, R. J., Noterdaeme, P., Johnson, J. W., et al. 2022, ApJ, 932,
    60, doi: 10.3847/1538-4357/ac6503

Curti, M., D'Eugenio, F., Carniani, S., et al. 2023, MNRAS, 518,
    425, doi: 10.1093/mnras/stac2737

Curti, M., Maiolino, R., Curtis-Lake, E., et al. 2024, A&A, 684,
    A75, doi: 10.1051/0004-6361/202346698

da Silva, R., D'Orazi, V., Palla, M., et al. 2023, A&A, 678, A195,
    doi: 10.1051/0004-6361/202346982

Daflon, S., & Cunha, K. 2004, ApJ, 617, 1115,
    doi: 10.1086/425607

Dahmen, G., Huttemeister, S., Wilson, T. L., & Mauersberger, R.
    1998, A&A, 331, 959, doi: 10.48550/arXiv.astro-ph/9711117

Dalcanton, J. J. 2007, ApJ, 658, 941, doi: 10.1086/508913

De Ridder, J., Barban, C., Baudin, F., et al. 2009, Nature, 459, 398,
    doi: 10.1038/nature08022

De Silva, G. M., Sneden, C., Paulson, D. B., et al. 2006, AJ, 131,
    455, doi: 10.1086/497968

De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015,
    MNRAS, 449, 2604, doi: 10.1093/mnras/stv327

Di Teodoro, E. M., & Peek, J. E. G. 2021, ApJ, 923, 220,
    doi: 10.3847/1538-4357/ac2cbd

Donnan, C. T., McLure, R. J., Dunlop, J. S., et al. 2024, arXiv
    e-prints, arXiv:2403.03171, doi: 10.48550/arXiv.2403.03171

Dubay, L. O., Johnson, J. A., & Johnson, J. W. 2024, arXiv
    e-prints, arXiv:2404.08059, doi: 10.48550/arXiv.2404.08059

Eilers, A.-C., Hogg, D. W., Rix, H.-W., et al. 2022, ApJ, 928, 23,
    doi: 10.3847/1538-4357/ac54ad

Ertl, T., Janka, H. T., Woosley, S. E., Sukhbold, T., & Ugliano, M.
    2016, ApJ, 818, 124, doi: 10.3847/0004-637X/818/2/124

Farmer, R., Laplace, E., de Mink, S. E., & Justham, S. 2021, ApJ,
    923, 214, doi: 10.3847/1538-4357/ac2f44

Ferrara, A., Pallottini, A., & Dayal, P. 2023, MNRAS, 522, 3986,
    doi: 10.1093/mnras/stad1095

Ferreira, L., Adams, N., Conselice, C. J., et al. 2022, ApJL, 938,
    L2, doi: 10.3847/2041-8213/ac947c

Ferreira, L., Conselice, C. J., Sazonova, E., et al. 2023, ApJ, 955,
    94, doi: 10.3847/1538-4357/acec76

Feuillet, D. K., Frankel, N., Lind, K., et al. 2019, MNRAS, 489,
    1742, doi: 10.1093/mnras/stz2221

Feuillet, D. K., Bovy, J., Holtzman, J., et al. 2018, MNRAS, 477,
    2326, doi: 10.1093/mnras/sty779

Finlator, K., & Davé, R. 2008, MNRAS, 385, 2181,
    doi: 10.1111/j.1365-2966.2008.12991.x

Frankel, N., Rix, H.-W., Ting, Y.-S., Ness, M., & Hogg, D. W.
    2018, ApJ, 865, 96, doi: 10.3847/1538-4357/aadba5

Friel, E. D. 1995, ARA&A, 33, 381,
    doi: 10.1146/annurev.aa.33.090195.002121

Frinchaboy, P. M., Thompson, B., Jackson, K. M., et al. 2013,
    ApJL, 777, L1, doi: 10.1088/2041-8205/777/1/L1

Frischknecht, U., Hirschi, R., Pignatari, M., et al. 2016, MNRAS,
    456, 1803, doi: 10.1093/mnras/stv2723

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021,
    A&A, 649, A1, doi: 10.1051/0004-6361/202039657

Gallart, C., Surot, F., Cassisi, S., et al. 2024, arXiv e-prints,
    arXiv:2402.09399, doi: 10.48550/arXiv.2402.09399

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., &
    Tremonti, C. A. 2005, MNRAS, 362, 41,
    doi: 10.1111/j.1365-2966.2005.09321.x

Gilmore, G., Randich, S., Asplund, M., et al. 2012, The Messenger,
    147, 25

Girardi, L. 2016, ARA&A, 54, 95,
    doi: 10.1146/annurev-astro-081915-023354

Gjergo, E., Sorokin, A. G., Ruth, A., et al. 2023, ApJS, 264, 44, doi: 10.3847/1538-4365/aca7c7

Grand, R. J. J., Kawata, D., & Cropper, M. 2014, MNRAS, 439, 623, doi: 10.1093/mnras/stt2483

Grand, R. J. J., Springel, V., Kawata, D., et al. 2016, MNRAS, 460, L94, doi: 10.1093/mnrasl/slw086

Griffith, E. J., Sukhbold, T., Weinberg, D. H., et al. 2021, ApJ, 921, 73, doi: 10.3847/1538-4357/ac1bac

Grisoni, V., Spitoni, E., & Matteucci, F. 2018, MNRAS, 481, 2570, doi: 10.1093/mnras/sty2444

Gutcke, T. A., Stinson, G. S., Macciò, A. V., Wang, L., & Dutton, A. A. 2017, MNRAS, 464, 2796, doi: 10.1093/mnras/stw2539

Hatchfield, H. P., Battersby, C., Keto, E., et al. 2020, ApJS, 251, 14, doi: 10.3847/1538-4365/abb610

Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, ApJ, 808, 132, doi: 10.1088/0004-637X/808/2/132

Hekker, S., & Christensen-Dalsgaard, J. 2017, A&A Rv, 25, 1, doi: 10.1007/s00159-017-0101-x

Henry, R. B. C., Edmunds, M. G., & Köppen, J. 2000, ApJ, 541, 660, doi: 10.1086/309471

Henry, R. B. C., Kwitter, K. B., Jaskot, A. E., et al. 2010, ApJ, 724, 748, doi: 10.1088/0004-637X/724/1/748

Hu, C.-Y., Smith, M. C., Teyssier, R., et al. 2022, arXiv e-prints, arXiv:2208.10528, doi: 10.48550/arXiv.2208.10528

Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, MNRAS, 315, 543, doi: 10.1046/j.1365-8711.2000.03426.x

Imig, J., Price, C., Holtzman, J. A., et al. 2023, ApJ, 954, 124, doi: 10.3847/1538-4357/ace9b8

Johnson, J. A. 2019, Science, 363, 474, doi: 10.1126/science.aau9540

Johnson, J. W., Kochanek, C. S., & Stanek, K. Z. 2023a, MNRAS, 526, 5911, doi: 10.1093/mnras/stad3019

Johnson, J. W., & Weinberg, D. H. 2020, MNRAS, 498, 1364, doi: 10.1093/mnras/staa2431

Johnson, J. W., Weinberg, D. H., Vincenzo, F., Bird, J. C., & Griffith, E. J. 2023b, MNRAS, 520, 782, doi: 10.1093/mnras/stad057

Johnson, J. W., Weinberg, D. H., Vincenzo, F., et al. 2021, MNRAS, 508, 4484, doi: 10.1093/mnras/stab2718

Johnson, J. W., Conroy, C., Johnson, B. D., et al. 2023c, MNRAS, 526, 5084, doi: 10.1093/mnras/stad2985

Jönsson, H., Allende Prieto, C., Holtzman, J. A., et al. 2018, AJ, 156, 126, doi: 10.3847/1538-3881/aad4f5

Kalberla, P. M. W., & Kerp, J. 2009, ARA&A, 47, 27, doi: 10.1146/annurev-astro-082708-101823

Kauffmann, G. 1996, MNRAS, 281, 475, doi: 10.1093/mnras/281.2.475

Kennicutt, Robert C., J. 1998, ApJ, 498, 541, doi: 10.1086/305588

Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102, doi: 10.1088/0004-637X/779/2/102

Kodama, T., & Arimoto, N. 1997, A&A, 320, 41, doi: 10.48550/arXiv.astro-ph/9609160

Kopenhafer, C., O'Shea, B. W., & Voit, G. M. 2023, ApJ, 951, 107, doi: 10.3847/1538-4357/accbb7

Kordopatis, G., Schultheis, M., McMillan, P. J., et al. 2023, A&A, 669, A104, doi: 10.1051/0004-6361/202244283

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Krumholz, M. R., Burkhart, B., Forbes, J. C., & Crocker, R. M. 2018, MNRAS, 477, 2716, doi: 10.1093/mnras/sty852

Kubryk, M., Prantzos, N., & Athanassoula, E. 2013, MNRAS, 436, 1479, doi: 10.1093/mnras/stt1667

Labbé, I., van Dokkum, P., Nelson, E., et al. 2023, Nature, 616, 266, doi: 10.1038/s41586-023-05786-2

Lacey, C. G., & Fall, S. M. 1985, ApJ, 290, 154, doi: 10.1086/162970

Larson, R. B. 1974, MNRAS, 166, 585, doi: 10.1093/mnras/166.3.585

Law, D. R., & Majewski, S. R. 2010, ApJ, 714, 229, doi: 10.1088/0004-637X/714/1/229

LeCun, Y., Bengio, Y., & Hinton, G. 2015, Nature, 521, 436, doi: 10.1038/nature14539

Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782, doi: 10.1088/0004-6256/136/6/2782

Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, AJ, 146, 19, doi: 10.1088/0004-6256/146/2/19

Leung, H. W., & Bovy, J. 2019a, MNRAS, 483, 3255, doi: 10.1093/mnras/sty3217

—. 2019b, MNRAS, 489, 2079, doi: 10.1093/mnras/stz2245

Leung, H. W., Bovy, J., Mackereth, J. T., & Miglio, A. 2023, MNRAS, doi: 10.1093/mnras/stad1272

Li, H., Vogelsberger, M., Marinacci, F., Sales, L. V., & Torrey, P. 2020, MNRAS, 499, 5862, doi: 10.1093/mnras/staa3122

Lian, J., Thomas, D., Maraston, C., et al. 2020a, MNRAS, 494, 2561, doi: 10.1093/mnras/staa867

—. 2020b, MNRAS, 497, 2371, doi: 10.1093/mnras/staa2078

Licquia, T. C., & Newman, J. A. 2015, ApJ, 806, 96, doi: 10.1088/0004-637X/806/1/96

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119, doi: 10.1088/0004-637X/772/2/119

Liu, F., Asplund, M., Yong, D., et al. 2016, MNRAS, 463, 696, doi: 10.1093/mnras/stw2045

Loebman, S. R., Roškar, R., Debattista, V. P., et al. 2011, ApJ, 737, 8, doi: 10.1088/0004-637X/737/1/8

Lopez, L. A., Mathur, S., Nguyen, D. D., Thompson, T. A., & Olivier, G. M. 2020, ApJ, 904, 152, doi: 10.3847/1538-4357/abc010

Lopez, S., Lopez, L. A., Nguyen, D. D., et al. 2023, ApJ, 942, 108, doi: 10.3847/1538-4357/aca65e

Lu, Y., Minchev, I., Buck, T., et al. 2022, arXiv e-prints,
    arXiv:2212.04515, doi: 10.48550/arXiv.2212.04515

Luck, R. E., Andrievsky, S. M., Kovtyukh, V. V., Gieren, W., &
    Graczyk, D. 2011, AJ, 142, 51,
    doi: 10.1088/0004-6256/142/2/51

Luck, R. E., Kovtyukh, V. V., & Andrievsky, S. M. 2006, AJ, 132,
    902, doi: 10.1086/505687

Luck, R. E., & Lambert, D. L. 2011, AJ, 142, 136,
    doi: 10.1088/0004-6256/142/4/136

Luo, A. L., Zhao, Y.-H., Zhao, G., et al. 2015, Research in
    Astronomy and Astrophysics, 15, 1095,
    doi: 10.1088/1674-4527/15/8/002

Maciel, W. J., Costa, R. D. D., & Uchida, M. M. M. 2003, A&A,
    397, 667, doi: 10.1051/0004-6361:20021530

Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, MNRAS, 489,
    176, doi: 10.1093/mnras/stz1521

Maeder, A., & Meynet, G. 1989, A&A, 210, 155

Magrini, L., Coccato, L., Stanghellini, L., Casasola, V., & Galli, D.
    2016, A&A, 588, A91, doi: 10.1051/0004-6361/201527799

Magrini, L., Sestito, P., Randich, S., & Galli, D. 2009, A&A, 494,
    95, doi: 10.1051/0004-6361:200810634

Maiolino, R., & Mannucci, F. 2019, A&A Rv, 27, 3,
    doi: 10.1007/s00159-018-0112-2

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ,
    154, 94, doi: 10.3847/1538-3881/aa784d

Martell, S. L., Sharma, S., Buder, S., et al. 2017, MNRAS, 465,
    3203, doi: 10.1093/mnras/stw2835

Matsuno, T., Aoki, W., Casagrande, L., et al. 2021, ApJ, 912, 72,
    doi: 10.3847/1538-4357/abeab2

Matteucci, F. 2021, A&A Rv, 29, 5,
    doi: 10.1007/s00159-021-00133-8

Matteucci, F., & Francois, P. 1989, MNRAS, 239, 885,
    doi: 10.1093/mnras/239.3.885

Melioli, C., Brighenti, F., D'Ercole, A., & de Gouveia Dal Pino,
    E. M. 2008, MNRAS, 388, 573,
    doi: 10.1111/j.1365-2966.2008.13446.x

—. 2009, MNRAS, 399, 1089,
    doi: 10.1111/j.1365-2966.2009.14725.x

Méndez-Delgado, J. E., Amayo, A., Arellano-Córdova, K. Z., et al.
    2022, MNRAS, 510, 4436, doi: 10.1093/mnras/stab3782

Miglio, A., Chiappini, C., Mackereth, J. T., et al. 2021, A&A, 645,
    A85, doi: 10.1051/0004-6361/202038307

Minchev, I., Chiappini, C., & Martig, M. 2013, A&A, 558, A9,
    doi: 10.1051/0004-6361/201220189

—. 2014, A&A, 572, A92, doi: 10.1051/0004-6361/201423487

Minchev, I., Famaey, B., Combes, F., et al. 2011, A&A, 527, A147,
    doi: 10.1051/0004-6361/201015139

Minchev, I., Anders, F., Recio-Blanco, A., et al. 2018, MNRAS,
    481, 1645, doi: 10.1093/mnras/sty2033

Moe, M., Kratter, K. M., & Badenes, C. 2019, ApJ, 875, 61,
    doi: 10.3847/1538-4357/ab0d88

Montalbán, J., Mackereth, J. T., Miglio, A., et al. 2021, Nature
    Astronomy, 5, 640, doi: 10.1038/s41550-021-01347-7

Morris, M., & Serabyn, E. 1996, ARA&A, 34, 645,
    doi: 10.1146/annurev.astro.34.1.645

Myers, N., Donor, J., Spoo, T., et al. 2022, AJ, 164, 85,
    doi: 10.3847/1538-3881/ac7ce5

Naidu, R. P., Oesch, P. A., van Dokkum, P., et al. 2022, ApJL, 940,
    L14, doi: 10.3847/2041-8213/ac9b22

Nelson, D., Pillepich, A., Springel, V., et al. 2019, MNRAS, 490,
    3234, doi: 10.1093/mnras/stz2306

Ness, M., Hogg, D. W., Rix, H. W., et al. 2016, ApJ, 823, 114,
    doi: 10.3847/0004-637X/823/2/114

Padovani, P., & Matteucci, F. 1993, ApJ, 416, 26,
    doi: 10.1086/173212

Palla, M., Matteucci, F., Spitoni, E., Vincenzo, F., & Grisoni, V.
    2020, MNRAS, 498, 1710, doi: 10.1093/mnras/staa2437

Palla, M., Santos-Peral, P., Recio-Blanco, A., & Matteucci, F.
    2022, A&A, 663, A125, doi: 10.1051/0004-6361/202142645

Peeples, M. S., & Shankar, F. 2011, MNRAS, 417, 2962,
    doi: 10.1111/j.1365-2966.2011.19456.x

Pejcha, O., & Thompson, T. A. 2015, ApJ, 801, 90,
    doi: 10.1088/0004-637X/801/2/90

Pessi, T., Anderson, J. P., Lyman, J. D., et al. 2023, ApJL, 955,
    L29, doi: 10.3847/2041-8213/acf7c6

Pierce-Price, D., Richer, J. S., Greaves, J. S., et al. 2000, ApJL,
    545, L121, doi: 10.1086/317884

Pinsonneault, M. H., Elsworth, Y., Epstein, C., et al. 2014, ApJS,
    215, 19, doi: 10.1088/0067-0049/215/2/19

Pinsonneault, M. H., Elsworth, Y. P., Tayar, J., et al. 2018, ApJS,
    239, 32, doi: 10.3847/1538-4365/aaebfd

Portinari, L., & Chiosi, C. 2000, A&A, 355, 929,
    doi: 10.48550/arXiv.astro-ph/0002145

Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2023, A&A,
    673, A155, doi: 10.1051/0004-6361/202245399

Ratcliffe, B., Minchev, I., Anders, F., et al. 2023, MNRAS, 525,
    2208, doi: 10.1093/mnras/stad1573

Rhoads, J. E., Wold, I. G. B., Harish, S., et al. 2023, ApJL, 942,
    L14, doi: 10.3847/2041-8213/acaaaf

Rodríguez, Ó., Maoz, D., & Nakar, E. 2022, arXiv e-prints,
    arXiv:2209.05552, doi: 10.48550/arXiv.2209.05552

Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., &
    Wadsley, J. 2008a, ApJL, 684, L79, doi: 10.1086/592231

Roškar, R., Debattista, V. P., Stinson, G. S., et al. 2008b, ApJL,
    675, L65, doi: 10.1086/586734

Ruiz-Lara, T., Gallart, C., Bernard, E. J., & Cassisi, S. 2020,
    Nature Astronomy, 4, 965, doi: 10.1038/s41550-020-1097-0

Sánchez, S. F. 2020, ARA&A, 58, 99,
    doi: 10.1146/annurev-astro-012120-013326

Sanders, R. L., Shapley, A. E., Jones, T., et al. 2021, ApJ, 914, 19, doi: 10.3847/1538-4357/abf4c1

Sandford, N. R., Weinberg, D. H., Weisz, D. R., & Fu, S. W. 2024, MNRAS, 530, 2315, doi: 10.1093/mnras/stae1010

Santana, F. A., Beaton, R. L., Covey, K. R., et al. 2021, AJ, 162, 303, doi: 10.3847/1538-3881/ac2cbc

Schaerer, D., Marques-Chaves, R., Barrufet, L., et al. 2022, A&A, 665, L4, doi: 10.1051/0004-6361/202244556

Schönrich, R., & Binney, J. 2009, MNRAS, 396, 203, doi: 10.1111/j.1365-2966.2009.14750.x

Searle, L. 1971, ApJ, 168, 327, doi: 10.1086/151090

Sellwood, J. A., & Binney, J. J. 2002, MNRAS, 336, 785, doi: 10.1046/j.1365-8711.2002.05806.x

Sharda, P., Krumholz, M. R., Wisnioski, E., et al. 2021, MNRAS, 502, 5935, doi: 10.1093/mnras/stab252

Shaver, P. A., McGee, R. X., Newton, L. M., Danks, A. C., & Pottasch, S. R. 1983, MNRAS, 204, 53, doi: 10.1093/mnras/204.1.53

Silva Aguirre, V., Bojsen-Hansen, M., Slumstrup, D., et al. 2018, MNRAS, 475, 5487, doi: 10.1093/mnras/sty150

Simon, J. D. 2019, ARA&A, 57, 375, doi: 10.1146/annurev-astro-091918-104453

Soderblom, D. R. 2010, ARA&A, 48, 581, doi: 10.1146/annurev-astro-081309-130806

Spina, L., Magrini, L., & Cunha, K. 2022, Universe, 8, 87, doi: 10.3390/universe8020087

Spitoni, E., & Matteucci, F. 2011, A&A, 531, A72, doi: 10.1051/0004-6361/201015749

Spitoni, E., Matteucci, F., Recchi, S., Cescutti, G., & Pipino, A. 2009, A&A, 504, 87, doi: 10.1051/0004-6361/200911768

Spitoni, E., Recchi, S., & Matteucci, F. 2008, A&A, 484, 743, doi: 10.1051/0004-6361:200809403

Spitoni, E., Silva Aguirre, V., Matteucci, F., Calura, F., & Grisoni, V. 2019, A&A, 623, A60, doi: 10.1051/0004-6361/201834188

Spitoni, E., Verma, K., Silva Aguirre, V., et al. 2021, A&A, 647, A73, doi: 10.1051/0004-6361/202039864

Spitzer, Lyman, J. 1942, ApJ, 95, 329, doi: 10.1086/144407

Stanghellini, L., & Haywood, M. 2010, ApJ, 714, 1096, doi: 10.1088/0004-637X/714/2/1096

Stone-Martinez, A., Holtzman, J. A., Imig, J., et al. 2024, AJ, 167, 73, doi: 10.3847/1538-3881/ad12a6

Sukhbold, T., Ertl, T., Woosley, S. E., Brown, J. M., & Janka, H. T. 2016, ApJ, 821, 38, doi: 10.3847/0004-637X/821/1/38

Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179, doi: 10.3847/1538-4357/aaa4b4

Tinsley, B. M. 1980, FCPh, 5, 287, doi: 10.48550/arXiv.2203.02041

Trapp, C. W., Kereš, D., Chan, T. K., et al. 2022, MNRAS, 509, 4149, doi: 10.1093/mnras/stab3251

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898, doi: 10.1086/423264

Trump, J. R., Arrabal Haro, P., Simons, R. C., et al. 2023, ApJ, 945, 35, doi: 10.3847/1538-4357/acba8a

Veilleux, S., Maiolino, R., Bolatto, A. D., & Aalto, S. 2020, A&A Rv, 28, 2, doi: 10.1007/s00159-019-0121-9

Vincenzo, F., & Kobayashi, C. 2020, MNRAS, 496, 80, doi: 10.1093/mnras/staa1451

Vincenzo, F., Weinberg, D. H., Miglio, A., Lane, R. R., & Roman-Lopes, A. 2021, MNRAS, 508, 5903, doi: 10.1093/mnras/stab2899

Weinberg, D. H., Andrews, B. H., & Freudenburg, J. 2017, ApJ, 837, 183, doi: 10.3847/1538-4357/837/2/183

Weinberg, D. H., Griffith, E. J., Johnson, J. W., & Thompson, T. A. 2023, arXiv e-prints, arXiv:2309.05719, doi: 10.48550/arXiv.2309.05719

Weinberg, D. H., Holtzman, J. A., Hasselquist, S., et al. 2019, ApJ, 874, 102, doi: 10.3847/1538-4357/ab07c7

Wenger, T. V., Balser, D. S., Anderson, L. D., & Bania, T. M. 2019, ApJ, 887, 114, doi: 10.3847/1538-4357/ab53d3

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52, doi: 10.1086/170483

Willett, E., Miglio, A., Mackereth, J. T., et al. 2023, MNRAS, 526, 2141, doi: 10.1093/mnras/stad2374

Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001, doi: 10.1088/1538-3873/ab0075

Wiseman, P., Sullivan, M., Smith, M., et al. 2021, MNRAS, 506, 3330, doi: 10.1093/mnras/stab1943

Wyse, R. F. G., & Silk, J. 1989, ApJ, 339, 700, doi: 10.1086/167329

Xiang, M., & Rix, H.-W. 2022, Nature, 603, 599, doi: 10.1038/s41586-022-04496-5

Yong, D., Carney, B. W., Teixera de Almeida, M. L., & Pohl, B. L. 2006, AJ, 131, 2256, doi: 10.1086/500538

Zahid, H. J., Bresolin, F., Kewley, L. J., Coil, A. L., & Davé, R. 2012, ApJ, 750, 120, doi: 10.1088/0004-637X/750/2/120

Zaritsky, D. 1992, ApJL, 390, L73, doi: 10.1086/186375

Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, AJ, 146, 81, doi: 10.1088/0004-6256/146/4/81

Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, AJ, 154, 198, doi: 10.3847/1538-3881/aa8df9