The Astrophysical Journal, 973:55 (26pp), 2024 September 20 https://doi.org/10.3847/1538-4357/ad61df
© 2024. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Galactic Chemical Evolution Models Favor an Extended Type Ia Supernova Delay-time Distribution

Liam O. Dubay[1,2], Jennifer A. Johnson[1,2], and James W. Johnson[1,2,3]
[1] Department of Astronomy, The Ohio State University, 140 W. 18th Ave, Columbus, OH 43210, USA; dubay.11@osu.edu
[2] Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA
[3] Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101, USA
Received 2024 April 11; revised 2024 July 8; accepted 2024 July 9; published 2024 September 18

Abstract

Type Ia supernovae (SNe Ia) produce most of the Fe-peak elements in the Universe and therefore are a crucial ingredient in galactic chemical evolution models. SNe Ia do not explode immediately after star formation, and the delay-time distribution (DTD) has not been definitively determined by supernova surveys or theoretical models. Because the DTD also affects the relationship among age, [Fe/H], and [α/Fe] in chemical evolution models, comparison with observations of stars in the Milky Way is an important consistency check for any proposed DTD. We implement several popular forms of the DTD in combination with multiple star formation histories for the Milky Way in multizone chemical evolution models that include radial stellar migration. We compare our predicted interstellar medium abundance tracks, stellar abundance distributions, and stellar age distributions to the final data release of the Apache Point Observatory Galactic Evolution Experiment. We find that the DTD has the largest effect on the [α/Fe] distribution: a DTD with more prompt SNe Ia produces a stellar abundance distribution that is skewed toward a lower [α/Fe] ratio. While the DTD alone cannot explain the observed bimodality in the [α/Fe] distribution, in combination with an appropriate star formation history it affects the goodness of fit between the predicted and observed high-α sequence. Our model results favor an extended DTD with fewer prompt SNe Ia than the fiducial t^{-1} power law.

Unified Astronomy Thesaurus concepts: Galaxy chemical evolution (580); Chemical abundances (224); Chemical enrichment (225); the Milky Way (1054); Milky Way disk (1050); Milky Way evolution (1052)

1. Introduction

Galactic chemical evolution (GCE) studies seek to explain the observed distribution of metals throughout the Milky Way. Tinsley (1979) made a compelling case that the nonsolar [α/Fe][4] ratios seen by, e.g., Wallerstein (1962) were caused by the contributors of the Fe-peak elements having different stellar lifetimes than the contributors of the α-elements. Type Ia supernovae (SNe Ia), the thermonuclear explosions of carbon–oxygen white dwarfs (WDs), are responsible for a majority of the Fe produced in the Galaxy (Matteucci & Greggio 1986); meanwhile, core-collapse supernovae (CCSNe), the explosions of massive stars, produce the α-elements (e.g., O and Mg) in addition to a smaller fraction of Fe. SNe Ia are delayed by ∼0.04–10 Gyr after star formation events, as evidenced by observations in both star-forming and elliptical galaxies (e.g., Maza & van den Bergh 1976). This delayed enrichment leads to a decrease in [α/Fe] with increasing [Fe/H] (Matteucci & Greggio 1986). Therefore, the relative abundances of the α-elements and Fe as a function of stellar age trace the balance of SN rates over time.

The delay-time distribution (DTD) refers to the rate of SN Ia events per unit mass of star formation as a function of stellar population age (for a review, see Section 3.5 of Maoz et al. 2014). When the DTD is convolved with the Galactic star formation rate (SFR), it yields the overall SN Ia rate. The quantitative details of the relationship between [α/Fe] and [Fe/H] are set by the DTD, and as such it is a key parameter in GCE models. However, the DTD remains poorly constrained because it reflects the detailed evolution of the SN Ia progenitor systems, so different models for the progenitors of SNe Ia will naturally predict different forms for the DTD.

The explosion mechanism(s) of SNe Ia are not fully understood (for reviews, see Maoz et al. 2014; Livio & Mazzali 2018; Ruiter 2020; Liu et al. 2023). Two general production channels have been proposed. In the single-degenerate (SD) case, the WD accretes mass from a close nondegenerate companion until it surpasses ∼1.4 $M_{\odot}$ and explodes (Whelan & Iben 1973; Nomoto 1982; Yoon & Langer 2003). In the double-degenerate (DD) case, two WDs merge after a gravitational-wave inspiral (Iben & Tutukov 1984; Webbink 1984; Pakmor et al. 2012) or head-on collision (Benz et al. 1989; Thompson 2011). Searches for signs of interaction between the SN ejecta and a nondegenerate companion (e.g., Panagia et al. 2006; Chomiuk et al. 2016; Tucker et al. 2020; Fausnaugh et al. 2021; Dubay et al. 2022) or for a surviving companion (e.g., Schaefer & Pagnotta 2012; Do et al. 2021; Tucker & Shappee 2024) have placed tight constraints on the SD channel, heavily disfavoring it as the main pathway for producing "normal" SNe Ia. The DD channel is now the preferred model, but it faces issues with matching observed SN Ia rates because not all WD mergers necessarily lead to a thermonuclear explosion (e.g., Nomoto & Iben 1985; Saio & Nomoto 1998; Shen et al. 2012), and the progenitor systems are difficult to detect even within our own Galaxy (Rebassa-Mansergas et al. 2019).

[4] In standard bracket notation, $[X/Y] \equiv \log_{10}(X/Y) - \log_{10}(X/Y)_{\odot}$. In this paper, we will use [α/Fe] and [O/Fe] interchangeably, although observational studies will often use a combination of α-elements to calculate [α/Fe].

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

As a result of the uncertainty regarding SN Ia progenitors, theoretical models have yet to converge on a single prediction for the DTD. For the DD channel, assumptions about the distribution of WD separations and the rate of gravitational-wave inspiral suggest a broad $\sim t^{-1}$ DTD at long delay times ($\gtrsim$1 Gyr), but at short delays ($\lesssim$1 Gyr) the rate is limited by the need to produce two WDs (see Greggio 2005; Maoz et al. 2014). Triple or higher-order progenitor systems could also produce a t^{-1} DTD (Fang et al. 2018; Rajamuthukumar et al. 2023). The DTD that would result from the SD channel depends greatly on the assumptions of binary population synthesis, but in general is expected to cover a narrower range of delay times and may feature a steep exponential cutoff at the long end (e.g., Greggio 2005).

Surveys of SNe Ia can constrain the DTD by comparing the observed rate of SNe Ia to their host galaxy parameters (e.g., Mannucci et al. 2005; Heringer et al. 2019) or inferred star formation histories (SFHs; e.g., Maoz et al. 2012), measuring SN Ia rates in galaxy clusters (e.g., Maoz et al. 2010), or comparing the volumetric SN Ia rate to the cosmic SFH as a whole (e.g., Graur et al. 2014; Strolger et al. 2020). Early studies, which had limited sample sizes, produced unimodal (Strolger et al. 2004) or bimodal (Mannucci et al. 2006) DTDs where the majority of SNe Ia explode within a relatively narrow range of delay times. More recent studies have recovered broader DTD functions, with many converging on a declining power law of $\sim t^{-1}$ (e.g., Graur & Maoz 2013; Graur et al. 2015; Maoz & Graur 2017; Castrillo et al. 2021; Wiseman et al. 2021), though there is some evidence for a steeper slope in galaxy clusters (Maoz & Graur 2017; Friedmann & Maoz 2018). It is especially difficult to constrain the DTD for short delay times (Maoz & Mannucci 2012; Rodney et al. 2014) because of the need for SN Ia rates at long look-back times and uncertainties in the age estimates of stellar populations.

The uncertainties in the SN Ia DTD propagate into GCE models. In principle, the observed chemical abundance patterns should therefore contain information about the DTD, and by extension the progenitors of SNe Ia. The metallicity distribution function (MDF)[5] and distribution of [O/Fe] record the history of SN Ia enrichment as a function of stellar age and location in the Galaxy. A striking feature of the [α/Fe] distribution in the Milky Way disk is the distinct separation into two components, the high- and low-α sequences, at similar metallicity (e.g., Bensby et al. 2014). Since the [α/Fe] abundance reflects the ratio of CCSN to SN Ia enrichment, the DTD should influence the [α/Fe] bimodality.

A few studies have investigated different DTDs in one-zone chemical evolution models, but comparisons to abundance data have been limited to the solar neighborhood (e.g., Andrews et al. 2017; Palicio et al. 2023). Matteucci et al. (2009) compared five DTDs in a multizone GCE model and found that the agreement with observations worsens if the fraction of prompt ($t \lesssim 100$ Myr) SNe Ia is either too high or too low, but they were similarly limited by the available data for the solar neighborhood. Poulhazan et al. (2018) found that the prompt component of the DTD affects the peak and width of the [α/Fe] distribution in a cosmological smoothed-particle hydrodynamics simulation, but their simulation was not designed to reproduce the parameters of the Milky Way. The current era of large spectroscopic surveys such as the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017) and the ongoing Milky Way Mapper (Kollmeier et al. 2017) has made abundances across the Milky Way disk available for comparison to more sophisticated GCE models.

This paper presents a comprehensive look at the DTD in a multizone GCE model that can qualitatively reproduce the observed abundance structure of the Milky Way disk. A multizone approach allows for a radially dependent parameterization of the SFH, outflows, stellar migration, and abundance gradient that can better match observations across the Galactic disk. We evaluate a selection of DTDs from the literature with multiple SFHs and a prescription for radial stellar migration in the Versatile Integrator for Chemical Evolution (`VICE`; Johnson & Weinberg 2020). In Section 2, we present our models for the DTD and SFH and describe our observational sample. In Section 3, we detail our one-zone chemical evolution models and present results. In Section 4, we present the results of our multizone models and compare to observations. In Section 5, we discuss the implications for the DTD and future surveys. In Section 6, we summarize our conclusions.

2. Methods

We use `VICE` to run chemical evolution models that closely follow those of Johnson & Weinberg (2020) and Johnson et al. (2021, hereafter J21). We refer the interested reader to the former for details about the `VICE` package and to the latter for details about the model Milky Way disk, including the star formation law, radial density gradient, and outflows. Similar to J21, we adopt a prescription for radial migration based on the `h277` hydrodynamical simulation (Christensen et al. 2012). In Appendix C, we describe our method for determining the migration distance $\Delta R_{\rm gal}$ and midplane distance $|z|$ for each model stellar population. Our method produces smoother distributions in chemical abundance space than the simulation-based approach, but the abundance distributions are otherwise unaffected by this change. Table 1 summarizes our model parameters and the subsections in which we discuss them in detail.

2.1. Nucleosynthetic Yields

For simplicity and easier comparison to the results of J21, we focus our analysis on O and Fe, representing the α and Fe-peak elements, respectively. Both elements are produced by CCSNe. `VICE` adopts the instantaneous recycling approximation for CCSNe, so the equation that governs CCSN enrichment as a function of star formation for some element x is simply

$$\dot{M}_x^{\rm CC}(t) = y_x^{\rm CC}\dot{M}_\star(t), \tag{1}$$

where $y_x^{\rm CC}$ is the CCSN yield of element x per unit mass of star formation, and $\dot{M}_\star$ is the SFR. Following J21, who in turn adopt their CCSN yields from Chieffi & Limongi (2004) and Limongi & Chieffi (2006), we adopt $y_{\rm O}^{\rm CC} = 0.015$ and $y_{\rm Fe}^{\rm CC} = 0.0012$. The primary effect of these yields is to set the low-[Fe/H] "plateau" in [O/Fe], which represents pure CCSN enrichment. The chosen yields for this paper produce a plateau at [O/Fe] = 0.45; see Weinberg et al. (2023) for more discussion on the effect of the CCSN yields on chemical evolution.

[5] In this paper, we refer to the MDF and the distribution of [Fe/H] interchangeably.

Table 1
A Summary of Parameters and Their Fiducial Values for Our Chemical Evolution Models (See Discussion in Section 2)

Quantity	Fiducial Value(s)	Section	Description
$R_{\rm gal}$	[0, 20] kpc	4	Galactocentric radius
$\delta R_{\rm gal}$	100 pc	4	Width of each concentric ring
$\Delta R_{\rm gal}$	N/A	C	Change in orbital radius due to stellar migration
$p(\Delta R_{\rm gal}\|\tau, R_{\rm form})$	Equation (C1)	C	Probability density function of radial migration distance
z	[−3, 3] kpc	C	Distance from Galactic midplane at present day
$p(z\|\tau, R_{\rm final})$	Equation (C2)	C	Probability density function of Galactic midplane distance
Δt	10 Myr	4	Time-step size
$t_{\rm max}$	13.2 Gyr	4	Disk lifetime
n	8	4	Number of stellar populations formed per ring per time step
$R_{\rm SF}$	15.5 kpc	4	Maximum radius of star formation
$M_{g,0}$	0	2.3	Initial gas mass
$\dot{M}_r$	continuous	4	Recycling rate (Johnson & Weinberg 2020, Equation (2))
$R_{\rm Ia}(t)$	Equation (5)	2.2	Delay-time distribution of Type Ia supernovae
t_D	40 Myr	2.2	Minimum SN Ia delay time
$N_{\rm Ia}/M_\star$	$2.2\times10^{-3}\ M_\odot^{-1}$	2.1	SNe Ia per unit mass of stars formed (Maoz & Mannucci 2012)
$y_{\rm O}^{\rm CC}$	0.015	2.1	CCSN yield of O
$y_{\rm Fe}^{\rm CC}$	0.0012	2.1	CCSN yield of Fe
$y_{\rm O}^{\rm Ia}$	0	2.1	SN Ia yield of O
$y_{\rm Fe}^{\rm Ia}$	0.00214	2.1	SN Ia yield of Fe
$f_{\rm IO}(t\|R_{\rm gal})$	Equation (11)	2.3	Time dependence of the inside-out SFR
$f_{\rm LB}(t\|R_{\rm gal})$	Equation (12)	2.3	Time dependence of the late-burst SFR
$\tau_{\rm rise}$	2 Gyr	2.3	SFR rise timescale for inside-out and early-burst models
$\tau_{\rm EB}(t)$	Equation (13)	2.3	Time dependence of the early-burst SFE timescale
$f_{\rm EB}(t\|R_{\rm gal})$	Equation (14)	2.3	Time dependence of the early-burst infall rate
$f_{\rm TI}(t\|R_{\rm gal})$	Equation (16)	2.3	Time dependence of the two-infall infall rate
$\tau_\star$	2 Gyr	3	SFE timescale in one-zone models
$\eta(R_{\rm gal}=8\ {\rm kpc})$	2.15	3	Outflow mass-loading factor at the solar annulus
$\tau_{\rm sfh}(R_{\rm gal}=8\ {\rm kpc})$	15.1 Gyr	2.3	SFH timescale at the solar annulus

Note. We omit some parameters that are unchanged from J21; see their Table 1 for details.

Following the formalism of Weinberg et al. (2017), the rate of Fe contribution to the ISM from SNe Ia is

$$\dot{M}_{\rm Fe}^{\rm Ia}(t) = y_{\rm Fe}^{\rm Ia}\langle\dot{M}_\star\rangle_{\rm Ia}(t), \tag{2}$$

where $\langle\dot{M}_\star\rangle_{\rm Ia}(t)$ is the time-averaged SFR weighted by the DTD at time t and $y_{\rm Fe}^{\rm Ia}$ is the Fe yield of SNe Ia. Weinberg et al. (2017) show in their Appendix A that

$$\langle\dot{M}_\star\rangle_{\rm Ia} \equiv \frac{\int_0^t \dot{M}_\star(t')R_{\rm Ia}(t-t')dt'}{\int_{t_D}^{t_{\rm max}} R_{\rm Ia}(t')dt'}, \tag{3}$$

where $R_{\rm Ia}$ is the DTD in units of $M_\odot^{-1}\ {\rm yr}^{-1}$, t_D is the minimum SN Ia delay time, and $t_{\rm max}$ is the lifetime of the disk. The denominator of Equation (3) is therefore equal to $N_{\rm Ia}/M_\star$, the total number of SNe Ia per $M_\odot$ of stars formed.

The yield $y_{\rm Fe}^{\rm Ia}$ measures the mass of Fe produced by SNe Ia over the full duration of the DTD, which can be expressed as

$$y_{\rm Fe}^{\rm Ia} = m_{\rm Fe}^{\rm Ia}\int_{t_D}^{t_{\rm max}} R_{\rm Ia}(t')dt' = m_{\rm Fe}^{\rm Ia}\frac{N_{\rm Ia}}{M_\star}, \tag{4}$$

where $m_{\rm Fe}^{\rm Ia}$ is the average mass of Fe produced by a single SN Ia, and $N_{\rm Ia}/M_\star = 2.2 \pm 1 \times 10^{-3}\ M_\odot^{-1}$ is the average number of SNe Ia per mass of stars formed (Maoz & Mannucci 2012). Adjusting the value of $y_{\rm Fe}^{\rm Ia}$ primarily affects the end point of chemical evolution tracks in [O/Fe]–[Fe/H] space. Following J21, we adopt $y_{\rm Fe}^{\rm Ia} = 0.00214$. This yield is originally adapted from the W70 model of Iwamoto et al. (1999), but it is increased slightly so that the inside-out SFH produces stars with [O/Fe] ≈ 0.0 by the end of the model. The overall scale of the yields is inconsequential: a lower value of $y_{\rm Fe}^{\rm Ia}$ would produce similar results if compensated with a lower outflow mass-loading factor $\eta \equiv \dot{M}_{\rm out}/\dot{M}_\star$ (this is the yield–outflow degeneracy; see Weinberg et al. 2023; Sandford et al. 2024; and Appendix B of Johnson et al. 2023a). Palla (2021) studied the effect of different SN Ia yields on GCE models in detail.

2.2. Delay-time Distributions

We explore five different functional forms for the DTD: a two-population model, a single power law, an exponential, a broken power law with an initially flat plateau, and a model computed from triple-system dynamics. We also investigate one or two useful variations of the input parameters for each functional form. Figure 1 presents a selection of these DTDs, and Table 2 summarizes the parameters and median delay times ($t_{\rm med}$) for all of our DTDs. We use simple forms rather than simulated physical or analytic models of SNe Ia, for the sake of decreased computational time and easier interpretation of the model predictions. Physically motivated models of the DTD must contend with many unknown or poorly constrained parameters, so our simplified forms have the advantage of

Table 2
Summary of SN Ia DTDs Explored in This Paper (See Discussion in Section 2.2)

Model	Equation	Parameters	$t_{\rm med}$ (Gyr)	Similar to
Two-population	(6)	$t_{\rm max} = 0.05$ Gyr, $\sigma = 0.015$ Gyr, $\tau = 3$ Gyr	0.39	Mannucci et al. (2006)
Power-law	(7)	$\alpha = -1.4$	0.18	(Maoz & Graur 2017, cluster); Heringer et al. (2019)
Power-law	(7)	$\alpha = -1.1$	0.48	(Maoz & Graur 2017, field); Wiseman et al. (2021)
Exponential	(8)	$\tau = 1.5$ Gyr	1.08	(Greggio 2005, SD); Schönrich & Binney (2009); Weinberg et al. (2017)
Exponential	(8)	$\tau = 3$ Gyr	2.08	...
Plateau	(9)	$W = 0.3$ Gyr, $\alpha = -1.1$	1.08	(Greggio 2005, CLOSE DD)
Plateau	(9)	$W = 1$ Gyr, $\alpha = -1.1$	2.02	(Greggio 2005, WIDE DD)
Triple-system	(10)	$f_{\rm init} = 0.05 f_{\rm peak}$, $t_{\rm rise} = 0.5$ Gyr, $W = 0.5$ Gyr, $\alpha = -1.1$	2.50	Rajamuthukumar et al. (2023)

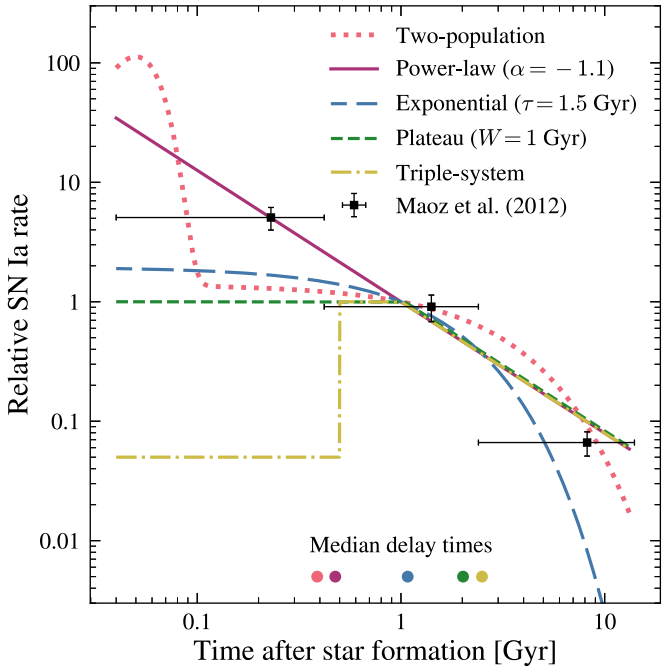

Figure 1. Selection of models for the SN Ia DTD used in this paper. All functions are normalized such that $f_{\rm Ia}(t = 1\ {\rm Gyr}) = 1$. The black squares represent the DTD recovered for the SDSS-II sample of SNe Ia by Maoz et al. (2012) at the same scale as the model DTDs. The horizontal and vertical error bars indicate the time range and 1σ uncertainties of each DTD measurement, respectively. The colored circles along the horizontal axis indicate the median delay time for each model.

reducing the number of free parameters. In Appendix B, we show that a few of our simple forms adequately approximate the more complete analytic models of Greggio (2005).

In this subsection, we present functional forms of each DTD in terms of a function $f_{\rm Ia}$ that has units of Gyr^{-1} and defines the shape of the DTD for $t \geqslant t_D$ as

$$R_{\rm Ia}(t) = \begin{cases} \dfrac{N_{\rm Ia}}{M_\star} \dfrac{f_{\rm Ia}(t)}{\int_{t_D}^{t_{\rm max}} f_{\rm Ia}(t')dt'}, & t \geqslant t_D \\ 0 & t < t_D. \end{cases} \quad (5)$$

Note that the denominator in Equation (5) normalizes the DTD.

Two-population. A DTD in which ~50% of SNe Ia belong to a "prompt" Gaussian component at small t and the remainder form an exponential tail at large t:

$$f_{\rm Ia}^{\rm twopop}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-t_p)^2}{2\sigma^2}} + \frac{1}{\tau} e^{-t/\tau}. \quad (6)$$

To approximate the DTD from Mannucci et al. (2006), we take $t_p = 50$ Myr, $\sigma = 15$ Myr, and $\tau = 3$ Gyr, which results in ~40% of SNe Ia exploding within $t < 100$ Myr. As we illustrate in Figure 1, the two-population DTD has a shorter median delay time than most other models (except the power-law model with $\alpha = -1.4$, not shown). This formulation is slightly different than the approximation used in other GCE studies (e.g., Matteucci et al. 2006; Poulhazan et al. 2018), where it has a more distinctly bimodal shape. We have compared the two approximations to this DTD in a one-zone model and found that they produce very similar abundance distributions. This DTD was adopted by the Feedback In Realistic Environments (FIRE; Hopkins et al. 2014) and FIRE-2 (Hopkins et al. 2018) simulations.

Power-law. A single power law with slope α:

$$f_{\rm Ia}^{\rm plaw}(t) = (t/1\ {\rm Gyr})^\alpha. \quad (7)$$

A declining power law with $\alpha \sim -1$ (Totani et al. 2008) arises from typical assumptions about the distribution of post-common-envelope separations and the rate of gravitational-wave inspiral (see Section 3.5 from Maoz et al. 2014). It is therefore a commonly assumed DTD in GCE studies (e.g., Rybizki et al. 2017; J21; Weinberg et al. 2023). Additionally, the observational evidence for a power-law DTD is strong. Maoz & Graur (2017) obtained a DTD with $\alpha = -1.07 \pm 0.09$ based on volumetric rates and an assumed cosmic SFH for field galaxies in redshift range $0 \leqslant z \leqslant 2.25$. Wiseman et al. (2021) obtained a similar slope of $\alpha = -1.13 \pm 0.05$ for field galaxies in the redshift range $0.2 < z < 0.6$. Heringer et al. (2019) used an SFH-independent method to constrain the DTD for field galaxies within $0.01 < z < 0.2$ and found a larger value of $\alpha = -1.34^{+0.19}_{-0.17}$. For galaxy clusters, Maoz & Graur (2017) found a steeper DTD slope of $\alpha = -1.39^{+0.32}_{-0.05}$, and Friedmann & Maoz (2018) found a similar slope of $\alpha = -1.3^{+0.23}_{-0.16}$; however, a reanalysis by Freundlich & Maoz (2021) revealed no significant difference between the cluster and field galaxy DTDs. In this paper, we investigate the cases $\alpha = -1.1$ and $\alpha = -1.4$.

Exponential. An exponentially declining DTD with timescale τ:

$$f_{\rm Ia}^{\rm exp}(t) = e^{-t/\tau}. \quad (8)$$

This model allows analytic solutions to the abundances as a function of time for some SFHs, making it a popular choice (e.g., Weinberg et al. 2017; Pantoni et al. 2019; Palicio et al. 2023). Schönrich & Binney (2009) and Weinberg et al. (2017) both assumed an exponential DTD with a timescale $\tau = 1.5$ Gyr. However, observational support for an exponential DTD is scarce. Strolger et al. (2020), fitting to the cosmic SFH and SFHs from field galaxies, found a range of exponential-like solutions with timescales $\sim$1.5–6 Gyr. In this paper, we investigate timescales $\tau = 1.5$ and 3 Gyr. We show in Appendix B that an exponential DTD with $\tau = 1.5$ Gyr is an adequate approximation for the analytic SD DTD from Greggio (2005).

Plateau. A modification of the power law in which the DTD "plateaus" for a duration W before declining:

$$f_{\rm Ia}^{\rm plat}(t) = \begin{cases} 1, & t < W \\ (t/W)^{\alpha}, & t \geqslant W. \end{cases} \quad (9)$$

Our primary motivation is to consider a model that matches observations at delay times beyond a few Gyr, where the DTD is best constrained, but with a smaller fraction of prompt ($\lesssim$100 Myr) SNe Ia than the single power law. This form has been used for the DTD of neutron star mergers (Simonetti et al. 2019), but to our knowledge it has not been considered for SNe Ia in previous GCE models. We show in Appendix B that this form can approximate the more complicated analytic DD DTDs from Greggio (2005). We investigate the cases $W = 0.3$ Gyr and $W = 1$ Gyr, taking $\alpha = -1.1$ for all plateau models.

Triple-system. A DTD based on simulations of triple-system evolution by Rajamuthukumar et al. (2023). We approximate their numerically generated DTD as a special case of the plateau model (Equation (9)) where the initial rate is quite low until an instantaneous rise to the plateau value at time $t_{\rm rise}$:

$$f_{\rm Ia}^{\rm triple}(t) = \begin{cases} \epsilon, & t < t_{\rm rise} \\ 1, & t_{\rm rise} \leqslant t < W \\ (t/W)^{\alpha}, & t \geqslant W, \end{cases} \quad (10)$$

with $t_{\rm rise} = 0.5$ Gyr, $W = 0.5$ Gyr, $\alpha = -1.1$, and $\epsilon = 0.05$ (i.e., the initial rate is 5% of the peak rate). As illustrated in Figure 1, the triple-system DTD has the longest median delay time out of all the models we investigate.

There are many models for the DTD that have been proposed in the literature, and an exhaustive test of every one is infeasible. In particular, we do not consider a Gaussian DTD (e.g., Strolger et al. 2004), in which most SNe Ia explode with delay times close to $\sim$3 Gyr and there are very few prompt events. Evidence for this form came from SN Ia rate measurements for $z > 1$, but reanalysis of the data revealed large errors due to the small sample size, extinction corrections, and uncertainties in the SFHs (e.g., Förster et al. 2006; Greggio et al. 2008). Additionally, Mannucci et al. (2006) found that such a DTD fails to reproduce the observed dependence of the SN Ia rate on galaxy color, and there is evidence that at least some SNe Ia must be prompt in order to explain observed rates in spiral galaxies (e.g., Mannucci et al. 2005; Scannapieco & Bildsten 2005). GCE studies have found that a Gaussian DTD overproduces high-[O/Fe] and low-[Fe/H] stars (Matteucci et al. 2009; Palicio et al. 2023).

2.2.1. The Minimum SN Ia Delay Time

In addition to the DTD shape, the minimum SN Ia delay time t_D is another parameter that can have an effect on chemical evolution observables, such as the location of the high-α knee and the [O/Fe] distribution function (DF; Andrews et al. 2017). The value of t_D is set by the lifetime of the most massive SN Ia progenitor system. Previous GCE studies have adopted values ranging from $t_D \approx 30$ Myr (e.g., Poulhazan et al. 2018) to $t_D = 150$ Myr (e.g., J21). We take $t_D = 40$ Myr as our fiducial value as it is the approximate lifetime of an $8\,M_\odot$ star. In Section 3, we find that adopting a longer t_D has only a minor effect on the chemical evolution for most DTDs except the power law, but in that case the effect of a longer t_D can be approximated by adding an initial plateau of width $W = 0.3$ Gyr to the DTD (see Figure 4).

2.3. Star Formation Histories

We consider four models for the SFH, which we refer to as inside-out, late-burst, early-burst, and two-infall. The former two models, which feature a smooth SFH, were investigated by J21 using a similar methodology to this paper. The inside-out model produced a good agreement to the age–[O/Fe] relation observed by Feuillet et al. (2019), while the late-burst model better matched their observed age–metallicity relation. The latter two models feature discontinuous or "bursty" SFHs. The early-burst model, proposed by Conroy et al. (2022), uses an efficiency-driven starburst to explain the break in the [α/Fe] trend observed in the H3 survey (Conroy et al. 2019). The two-infall model was proposed by Chiappini et al. (1997) and features two distinct episodes of gas infall, which produce the thick and thin disks. Together, these four models cover a range of behavior, including a smooth SFH, and SFR-, SFE-, and infall-driven starbursts.

The inside-out and late-burst models are run in `VICE`'s "star formation mode," where the SFR surface density $\dot{\Sigma}_\star$ is prescribed along with the star formation efficiency (SFE) timescale $\tau_\star \equiv \Sigma_g/\dot{\Sigma}_\star$. The remaining quantities, infall rate surface density $\dot{\Sigma}_{\rm in}$ and gas surface density Σ_g, are calculated from the specified quantities assuming the star formation law adopted by J21 (see their Equation (14)). More specifically, at each time step, the infall rate is calculated to fulfill the quantity of gas required to produce the specified SFR by the star formation law. The latter two models are run in "infall mode," where we specify $\dot{\Sigma}_{\rm in}$ and $\tau_\star$. The initial gas mass is zero for all models (including those run in star formation mode). The mode in which `VICE` models are run makes no difference, as a unique solution can always be obtained if two of the four parametric forms are specified.

The SFH is normalized such that the model predicts a total stellar mass of $(5.17 \pm 1.11) \times 10^{10} M_\odot$ (Licquia & Newman 2015) and the stellar surface density gradient reported by Bland-Hawthorn & Gerhard (2016) (see Appendix B of J21). We present an overview of the four SFHs in Figure 2, and we discuss them individually here.

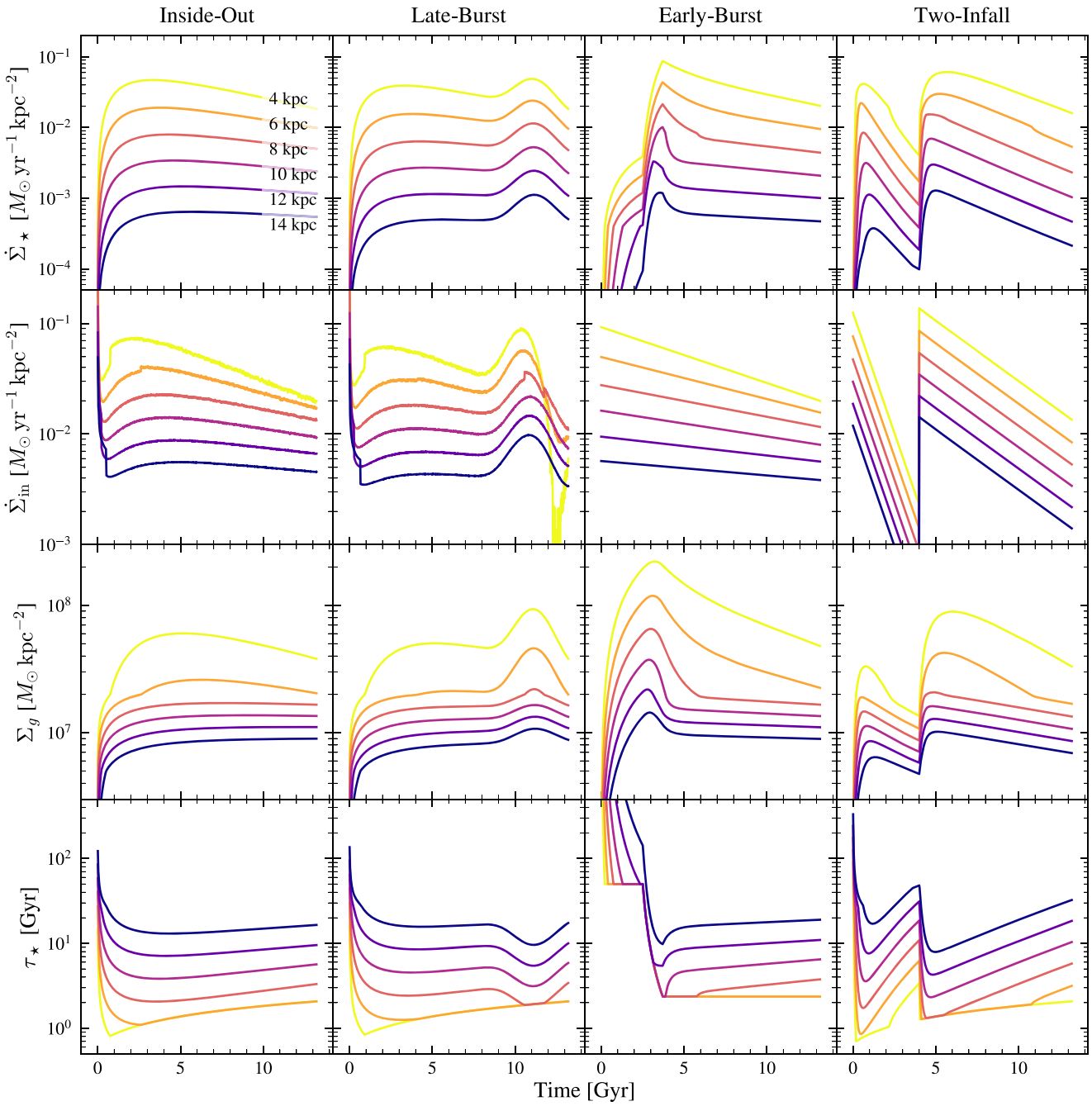

Figure 2. The surface densities of star formation $\dot{\Sigma}_\star$ (first row from the top), gas infall $\dot{\Sigma}_{\rm in}$ (second row), and gas mass Σ_g (third row), and the SFE timescale $\tau_\star$ (fourth row) as functions of time for our four model SFHs (see discussion in Section 2.3): inside-out (first column from the left; see Equation (11)), late-burst (second column; see Equation (12)), early-burst (third column; see Equations (13) and (14)), and two-infall (fourth column; see Equation (16)). In each panel, we plot curves for the model zones, which have inner radii at 4 kpc (yellow), 6 kpc (orange), 8 kpc (red), 10 kpc (violet), 12 kpc (indigo), and 14 kpc (blue).

Inside-out. As in J21, this is our fiducial SFH. The dimensionless time dependence of the SFR is given by

$$f_{\rm IO}(t|R_{\rm gal}) = \left[1 - \exp\left(\frac{-t}{\tau_{\rm rise}}\right)\right]\exp\left(\frac{-t}{\tau_{\rm sfh}(R_{\rm gal})}\right), \quad (11)$$

where we assume $\tau_{\rm rise} = 2$ Gyr for all radii. The SFH timescale $\tau_{\rm sfh}$ varies with $R_{\rm gal}$, with $\tau_{\rm sfh}(R_{\rm gal} = 8\,{\rm kpc}) \approx 15$ Gyr at the solar annulus and longer timescales in the outer Galaxy. The $\tau_{\rm sfh}$–$R_{\rm gal}$ relation is based on the radial gradients in stellar age in Milky Way–like spirals measured by Sánchez (2020); see Section 2.5 of J21 for details.

Late-burst. A variation on the inside-out SFH with a burst in the SFR at late times, which is described by a Gaussian according to

$$f_{\rm LB}(t|R_{\rm gal}) = f_{\rm IO}(t|R_{\rm gal})(1 + A_b e^{-(t-t_b)^2/2\sigma_b^2}), \quad (12)$$

where A_b is the dimensionless amplitude of the starburst, t_b is the time of the peak of the burst, and σ_b is the width of the Gaussian. Evidence for a recent star formation burst ∼2–3 Gyr ago has been found in Gaia (Mor et al. 2019) and in massive WDs in the solar neighborhood (Isern 2019). Following J21,

we adopt $A_b = 1.5$, $t_b = 11.2$ Gyr, and $\sigma_b = 1$ Gyr. The values of $\tau_{\rm rise}$ and $\tau_{\rm sfh}(R_{\rm gal})$ are the same as in the inside-out case.

Early-burst. An extension of the model proposed by Conroy et al. (2022) to explain the non-monotonic behavior of the high-α sequence down to [Fe/H] ≈ -2.5. This model features an abrupt factor ∼20 rise in the SFE at early times, driving an increase in the [O/Fe] abundance at the transition between the epochs of halo and thick-disk formation. Sahlholdt et al. (2022) found evidence for a burst ∼10 Gyr ago, which marks the beginning of a second phase of star formation. Mackereth et al. (2018) found that an early infall-driven burst of star formation can lead to an MW-like α-bimodality in the EAGLE simulations (Crain et al. 2015; Schaye et al. 2015). We adopt the following formula for the time dependence of the SFE timescale from Conroy et al. (2022):

$$\frac{\tau_{\rm EB}}{1\ {\rm Gyr}} = \begin{cases} 50, & t < 2.5\ {\rm Gyr} \\ \dfrac{50}{[1+3(t-2.5)]^2}, & 2.5 \leqslant t \leqslant 3.7\ {\rm Gyr} \\ 2.36, & t > 3.7\ {\rm Gyr}. \end{cases} \quad (13)$$

While Conroy et al. (2022) used a constant infall rate in their one-zone model, we adopt a radially dependent infall rate that declines exponentially with time:

$$f_{\rm EB}(t|R_{\rm gal}) = \exp\left(\frac{-t}{\tau_{\rm sfh}(R_{\rm gal})}\right), \quad (14)$$

where $\tau_{\rm sfh}$ is the same as in the inside-out case. To calculate $\dot{\Sigma}_\star$ from the above quantities, we modify the fiducial star formation law adopted from J21, substituting $\tau_{\rm EB}$ for the SFE timescale of molecular gas:

$$\dot{\Sigma}_\star = \begin{cases} \Sigma_g \tau_{\rm EB}^{-1}, & \Sigma_g \geqslant \Sigma_{g,2} \\ \Sigma_g \tau_{\rm EB}^{-1}\left(\dfrac{\Sigma_g}{\Sigma_{g,2}}\right)^{2.6}, & \Sigma_{g,1} \leqslant \Sigma_g \leqslant \Sigma_{g,2} \\ \Sigma_g \tau_{\rm EB}^{-1}\left(\dfrac{\Sigma_{g,1}}{\Sigma_{g,2}}\right)^{2.6}\left(\dfrac{\Sigma_g}{\Sigma_{g,1}}\right)^{0.7}, & \Sigma_g \leqslant \Sigma_{g,1}, \end{cases} \quad (15)$$

with $\Sigma_{g,1} = 5 \times 10^6\ M_\odot\ {\rm kpc}^{-2}$ and $\Sigma_{g,2} = 2 \times 10^7\ M_\odot\ {\rm kpc}^{-2}$.

Two-infall. First proposed by Chiappini et al. (1997), this model parameterizes the infall rate as two successive, exponentially declining bursts to explain the origin of the high- and low-α disk populations:

$$f_{\rm TI}(t|R_{\rm gal}) = N_1(R_{\rm gal})e^{-t/\tau_1} + N_2(R_{\rm gal})e^{-(t-t_{\rm on})/\tau_2}, \quad (16)$$

where $\tau_1 = 1$ Gyr and $\tau_2 = 4$ Gyr are the exponential timescales of the first and second infall, respectively, and $t_{\rm on} = 4$ Gyr is the onset time of the second infall (based on typical values in, e.g., Chiappini et al. 1997; Spitoni et al. 2020, 2021). N_1 and N_2 are the normalizations of the first and second infall, respectively, and their ratio N_2/N_1 is calculated so that the thick-to-thin-disk surface density ratio $f_\Sigma(R) = \Sigma_2(R)/\Sigma_1(R)$ is given by

$$f_\Sigma(R) = f_\Sigma(0)e^{R(1/R_2 - 1/R_1)}. \quad (17)$$

Following Bland-Hawthorn & Gerhard (2016), we adopt values for the thick-disk scale radius $R_1 = 2.0$ kpc, thin-disk scale radius $R_2 = 2.5$ kpc, and $f_\Sigma(0) = 0.27$. We note that most previous studies that use the two-infall model (e.g., Chiappini et al. 1997; Matteucci et al. 2006, 2009; Spitoni et al. 2019) do not consider gas outflows and instead adjust the nucleosynthetic yields to reproduce the solar abundance. We adopt radially dependent outflows as in J21 (see their Section 2.4 for details) for all our SFHs, including two-infall. We discuss the implications of this difference in Section 5.2.

Table 3
Sample Selection Parameters and Median Uncertainties from APOGEE DR17 (see Section 2.4)

Parameter	Range or Value	Notes
$\log g$	$1.0 < \log g < 3.8$	Select giants only
$T_{\rm eff}$	$3500 < T_{\rm eff} < 5500$ K	Reliable temperature range
S/N	S/N > 80	Required for accurate stellar parameters
ASPCAPFLAG Bits	$\notin 23$	Remove stars flagged as bad
EXTRATARG Bits	$\notin$ 0, 1, 2, 3, or 4	Select main red star sample only
Age	$\sigma_{\rm Age} < 40\%$	Age uncertainty from L23
$R_{\rm gal}$	$3 < R_{\rm gal} < 15$ kpc	Eliminate bulge and extreme outer-disk stars
$\lvert z\rvert$	$\lvert z\rvert < 2$ kpc	Eliminate halo stars

2.4. Observational Sample

We compare our model results to abundance measurements from the final data release (DR17; Abdurro'uf et al. 2022) of the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017). APOGEE used infrared spectrographs (Wilson et al. 2019) mounted on two telescopes: the 2.5 m Sloan Foundation Telescope (Gunn et al. 2006) at Apache Point Observatory in the Northern Hemisphere, and the Irénée DuPont Telescope (Bowen & Vaughan 1973) at Las Campanas Observatory in the Southern Hemisphere. After the spectra were passed through the data reduction pipeline (Nidever et al. 2015), the APOGEE Stellar Parameter and Chemical Abundance Pipeline (ASPCAP; Holtzman et al. 2015; García Pérez et al. 2016) extracted chemical abundances using the model grids and interpolation method described by Jönsson et al. (2020).

We restrict our sample to red giant branch and red clump stars with high-quality spectra. Table 3 lists our selection criteria, which largely follow from Hayden et al. (2015). This produces a final sample of 171,635 stars with calibrated [O/Fe] and [Fe/H] abundance measurements. We make use of the Gaia Early Data Release 3 (EDR3) data (Gaia Collaboration et al. 2016, 2021) included in the catalog by the APOGEE team. Specifically, we use the Bailer-Jones et al. (2021) photogeometric distance estimates to calculate Galactocentric radius $R_{\rm gal}$ and midplane distance z, assuming a Sun–Galactic center distance $R_\odot = 8.122$ kpc (GRAVITY Collaboration et al. 2018) and height of the Sun above the midplane $z_\odot = 20.8$ pc (Bennett & Bovy 2019). Table 4 lists the number of APOGEE stars in bins of $R_{\rm gal}$ and $|z|$. For some Galactic regions with $R_{\rm gal} < 5$ kpc or $R_{\rm gal} > 13$ kpc, the median distance error exceeds 1 kpc but remains within our bin width of 2 kpc, and the vast majority of stars have much smaller distance uncertainties.

We use estimated ages from Leung et al. (2023, hereafter L23), who use a variational encoder–decoder network that is trained on asteroseismic data to retrieve age estimates for APOGEE giants without contamination from age–abundance

Table 4
Number of APOGEE Stars in Each Galactic Region

$R_{gal} \in$ / $\lvert z\rvert \in$	(3, 5] kpc	(5, 7] kpc	(7, 9] kpc	(9, 11] kpc	(11, 13] kpc	(13, 15] kpc
(1.0, 2.0] kpc	2013	2100	8734	3663	1324	363
(0.5, 1.0] kpc	2487	3490	13,811	9069	3289	460
(0.0, 0.5] kpc	3296	7029	17,319	16,276	6336	812

correlations. Importantly, the L23 ages do not plateau beyond ∼10 Gyr as they do in astroNN (Mackereth et al. 2019). We use an age uncertainty cut of 40% per the recommendations of L23, which produces a total sample of 57,607 APOGEE stars with age estimates. We note that we use the full sample of 171,635 APOGEE stars unless we explicitly compare to age estimates. Table 5 presents the median and dispersion (95th − 5th percentile difference) of the uncertainty in [Fe/H], [O/Fe], and log(age).

Table 5
Median and Dispersion in APOGEE Parameter Uncertainties

Parameter	Median Uncertainty	Uncertainty Dispersion (95% − 5%)
[Fe/H]	0.0089	0.0060
[O/Fe]	0.019	0.031
log(Age/Gyr)	0.10	0.16

3. One-zone Models

Before running the full multizone models, it is useful to understand the effects of the DTD in more idealized conditions. A one-zone model assumes the entire gas reservoir is instantaneously mixed, removing all spatial dependence. This limits the ability to compare to observations across the disk, but it obviates the complicating factor of stellar migration and better isolates the effects of the nucleosynthesis prescription. In this section, we compare the results from one-zone models that examine various parameters of the DTD while keeping other parameters fixed. We use the outputs of our one-zone models to identify the regions in chemical abundance space that are most sensitive to the DTD.

For consistency, we adopt most of the parameter values from Table 1 for our one-zone models. We adopt the inside-out SFR (Equation (11)) evaluated at $R_{gal} = 8$ kpc (i.e., $\tau_{rise} = 2$ Gyr and $\tau_{sfh} = 15.1$ Gyr) and an SFE timescale $\tau_\star \equiv M_g/\dot{M}_\star = 2$ Gyr. Unless otherwise specified, we adopt an outflow mass-loading factor $\eta \equiv \dot{M}_{out}/\dot{M}_\star = 2.15$ (see Equation (8) from J21) and a minimum SN Ia delay time $t_D = 40$ Myr.

3.1. DTD Parameters: Slope, Timescale, and Plateau Width

The left-hand panel of Figure 3 compares the results of three one-zone models that are identical except for the slope of the power-law DTD. A steeper slope implies a greater number of prompt SNe Ia, which rapidly enrich the ISM with Fe, producing a faster decline in [O/Fe] with increasing [Fe/H] and hence a sharper "knee" after the minimum delay time. This results in a narrower distribution of [O/Fe] around the low-α sequence and a dearth of high-α stars. In all cases, the [O/Fe] DF is distinctly unimodal. The MDF is not as strongly affected by the power-law slope: a shallower slope results in only a modest increase in the width of the distribution. The abundance tracks converge to the equilibrium value, reflecting the yield ratio of CCSNe to SNe Ia, which is the same in all models.

Similar trends can be seen when adjusting the timescale of the exponential DTD, as shown in the middle panel of Figure 3. Here, the knee is not a sharp feature associated with the onset of SNe Ia as in the power-law case, but rather a gentle curve in the abundance track around $t = 1$ Gyr. Doubling the timescale from 1.5 Gyr to 3 Gyr implies a longer median delay time, which raises the [O/Fe] abundance ratio at $t = 1$ Gyr by ∼0.05 dex and at $t = 3$ Gyr by ∼0.1 dex. A longer exponential timescale also produces a broader [O/Fe] DF with more high-α stars, but the distribution is still unimodal. The effect on the MDF is slightly more pronounced than the power-law case, with longer timescales skewing to lower [Fe/H] values.

Finally, the right-hand panel of Figure 3 shows the effect of varying the width W of the plateau DTD. The abundance tracks from several different plateau widths fill the space in between the exponential ($\tau = 3$ Gyr) and power-law ($\alpha = -1.1$ with no plateau) models, which are both included in the panel for reference. The plateau ($W = 1$ Gyr) and exponential ($\tau = 3$ Gyr) DTDs produce nearly identical abundance tracks but their [O/Fe] DFs are more distinct, illustrating the need for both observables to discriminate between DTDs. The effect on the [O/Fe] DF is similar to the previous two models: a longer plateau raises the median delay time, producing a broader [O/Fe] DF and a more prominent high-α tail. On the other hand, all of the plateau DTDs produce very similar MDFs.

Unlike the previous three DTDs, we fix the hyperparameters of the two-population and triple-system models to reproduce specific DTDs from the literature. A variant of the two-population DTD with a broader Gaussian component ($t_{max} = 0.1$ Gyr, $\sigma = 0.03$ Gyr) produced similar abundance tracks and a nearly identical [O/Fe] distribution to the fiducial parameterization in a one-zone model. The effect of different parameters for the triple-system DTD would be very similar to the plateau DTD described above, of which the triple-system DTD is a special case.

3.2. The Minimum SN Ia Delay Time

We also explore the effect of varying the minimum SN Ia delay time t_D (Section 2.2.1). The left-hand panel of Figure 4 shows that t_D has a much stronger effect in models that assume a power-law DTD than others. This is a consequence of the high number of prompt SNe Ia ($t \lesssim 100$ Myr; see Figure 1). Moreover, a power-law DTD with a long t_D may be observationally hard to distinguish from a plateau model. In Figure 4, the abundance track for the model with a power-law DTD and $t_D = 150$ Myr (dashed purple line) is similar to that of the plateau DTD with $W = 0.3$ Gyr and $t_D = 40$ Myr (solid green line), and their [O/Fe] DFs are virtually identical. For the exponential ($\tau = 3$ Gyr) DTD, the two values of t_D produce nearly indistinguishable outputs. We do not consider the effect on the other DTDs because a 150 Myr minimum delay time is

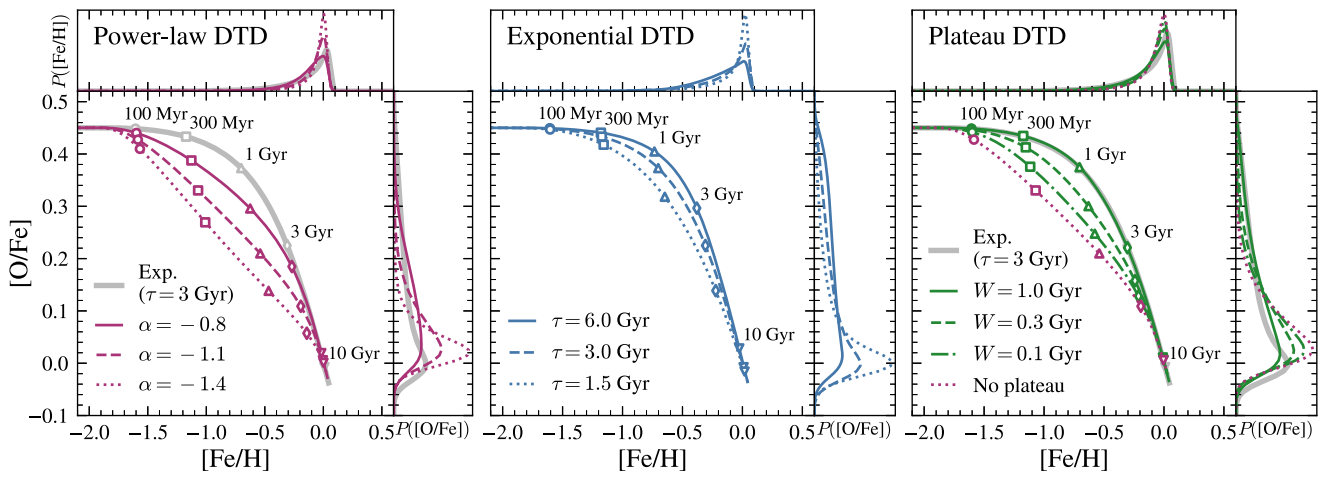

Figure 3. Abundance tracks in the [O/Fe]–[Fe/H] plane for one-zone chemical evolution models (see discussion in Section 3), which assume the various DTD shapes (see Figure 1). The open symbols along each curve mark logarithmic steps in time. The top and right-hand marginal panels present the distribution functions (DFs) of [Fe/H] and [O/Fe], respectively. For display purposes, these distributions are convolved with a Gaussian kernel with a standard deviation of 0.02 dex. Left: A power-law DTD with varying slope α. For reference, the solid gray curve represents an exponential DTD with $\tau = 3$ Gyr. Center: An exponential DTD with varying timescale τ. Right: A plateau DTD with varying width W. All assume a postplateau slope of $\alpha = -1.1$. For reference, the solid gray curve represents an exponential DTD with $\tau = 3$ Gyr, and the dotted purple curve represents a power-law DTD with $\alpha = -1.1$ and no plateau.

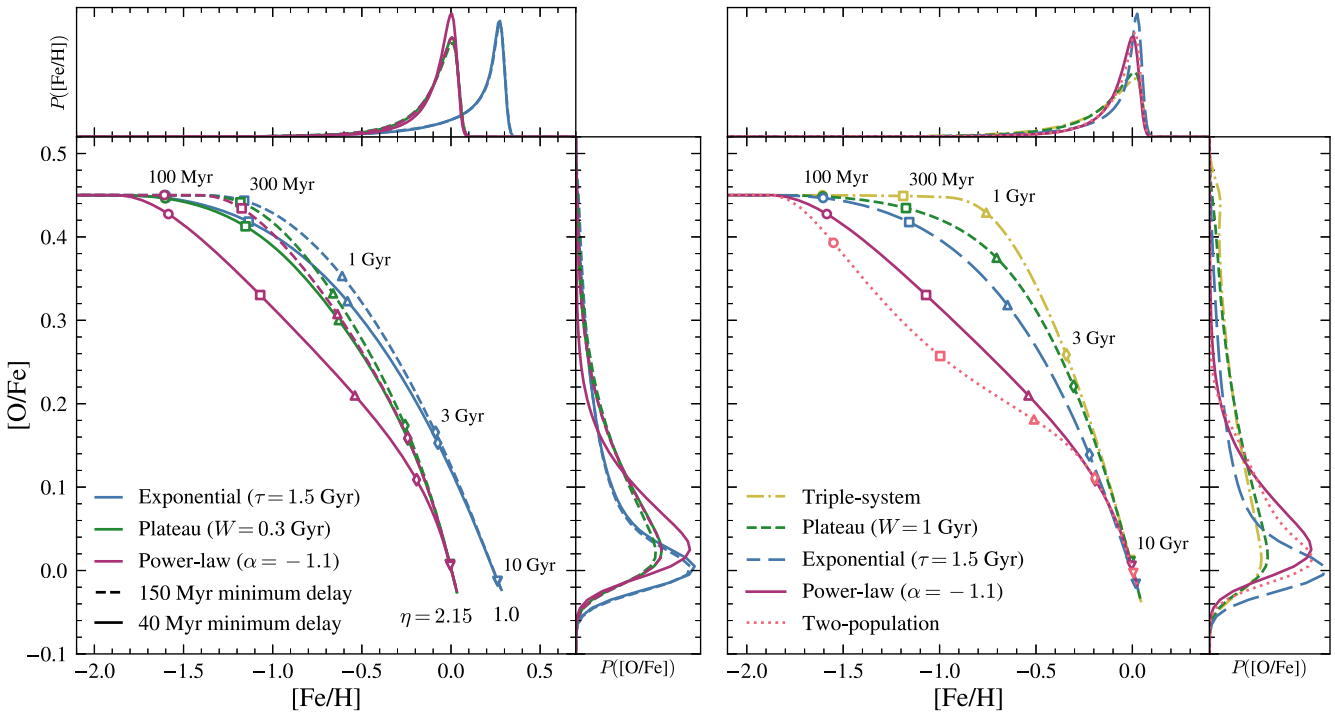

Figure 4. Left: Comparison of one-zone models with different combinations of minimum delay time t_D and DTD shape. The layout is similar to that of Figure 3. For visual clarity, we assume a mass-loading factor $\eta = 1$ for the exponential DTD curves, which places the end point of the abundance tracks at higher [Fe/H]. Right: comparison of one-zone models with five different DTD models (see Figure 1).

incompatible with the two-population model, which has ∼50% of SNe Ia explode in the first 100 Myr, and would have a negligible effect on the triple-system DTD due to its low SN Ia rate at short delay times. In the multizone models, we will hold t_D fixed at 40 Myr.

3.3. The Form of the DTD

The right-hand panel of Figure 4 compares the one-zone model outputs from the full range of DTDs we investigate in this paper. As with the individual DTD parameters, the form of the DTD primarily affects the location of the high-α knee in the [O/Fe]–[Fe/H] abundance tracks. At one extreme is the triple-system model, which sees the CCSN plateau extend up to [Fe/H] ≈ -0.8, followed by a sharp downward turn as the SN Ia rate suddenly increases at a delay time of 500 Myr. At the other extreme are the two-population and power-law ($\alpha = -1.1$) DTDs, for which the SN Ia rate peaks immediately after the minimum delay time of 40 Myr, placing the high-α knee at [Fe/H] ≈ -1.8. The two-population model has a unique second knee at [Fe/H] ≈ -0.2 and [O/Fe] ≈ 0.1, which is produced by the delayed exponential component, as noted by Vincenzo et al. (2017). The abundance tracks from the plateau

($W = 1$ Gyr) and exponential ($\tau = 1.5$ Gyr) models occupy the intermediate space between these extremes.

The [O/Fe] DFs also show significant differences between the DTDs. In the triple-system model, star formation proceeds for such a long time before the knee that the [O/Fe] DF shows a slight second peak around the CCSN yield ratio (∼0.45 dex). Out of all our one-zone models, this small bump is the only degree of bimodality that arises in the [O/Fe] DF. Below [O/Fe] ≈ 0.4, the plateau ($W = 1$ Gyr) and triple-system DTDs produce nearly identical distributions, while the exponential DTD produces the narrowest distribution. The power-law ($\alpha = -1.1$) and two-population DTDs produce similar [O/Fe] DFs despite notably different abundance tracks. The exponential ($\tau = 3$ Gyr) and plateau ($W = 0.3$ Gyr) models, while not shown, produce similar abundance tracks to the plateau ($W = 1$ Gyr) and exponential ($\tau = 1.5$ Gyr) models, respectively. The DTD also slightly shifts the peak of the [O/Fe] DF, with the exponential DTD placing it ∼0.02 dex lower than the power-law DTD. We see similar trends in the MDF, but to a lesser degree.

The results presented in this section indicate that the [O/Fe]–[Fe/H] abundance tracks and the [O/Fe] DF are most sensitive to the parameters of the DTD, while the MDF is a less sensitive diagnostic. Degeneracies between models in one regime can be resolved in the other. For example, the exponential ($\tau = 3$ Gyr) and plateau ($W = 1$ Gyr) DTDs are indistinguishable in [O/Fe]–[Fe/H] space but predict different [O/Fe] DFs. Of course, both of these observables are also greatly affected by the parameters of the SFH. In this section, we focused on the fiducial inside-out SFH. Palicio et al. (2023) compared similar DTDs in one-zone models with a two-infall SFH (see Section 5.2).

4. Multizone Models

We use the multizone GCE model tools in `VICE` developed by J21. The basic setup of our models follows theirs. The disk is divided into concentric rings of width $\delta R_{\rm gal} = 100$ pc. Stellar populations migrate radially under the prescription we describe in Appendix C, but each ring is otherwise described by a conventional one-zone GCE model with instantaneous mixing (see discussion in Section 3). Following J21, we do not implement radial gas flows (e.g., Lacey & Fall 1985; Bilitewski & Schönrich 2012). Stellar populations are also assigned a distance from the midplane according to their age and final radius as described in Appendix C.

We run our models with a time-step size of $\Delta t = 10$ Myr up to a maximum time of $t_{\rm max} = 13.2$ Gyr. Following J21, we set `VICE` to form $n = 8$ stellar populations per ring per time step, and we set a maximum star formation radius of $R_{\rm SF} = 15.5$ kpc, such that $\dot{\Sigma}_\star = 0$ for $R_{\rm gal} > R_{\rm SF}$. The model has a full radial extent of 20 kpc, allowing a purely migrated population to arise in the outer 4.5 kpc. We adopt continuous recycling, which accounts for the time-dependent return of mass from all previous generations of stars (see Equation (2) from Johnson & Weinberg 2020). We summarize these parameters in Table 1.

We run a total of multizone models with all combinations of our eight DTDs and four SFHs, for a total of 32. In the following subsections, we present the stellar abundance and age distributions from the multizone models and compare to APOGEE data from across the Galactic disk.

4.1. The Distribution of [Fe/H]

Figure 5 shows MDFs across the Galaxy for a selection of models and APOGEE data. The two left-hand columns illustrate the effect of different SFHs on the model outputs, which is most pronounced in the inner Galaxy. Near the midplane and in the inner Galaxy, the two-infall SFH produces a distinct bump ∼0.4 dex below the MDF peak, which is not seen for the inside-out SFH. Away from the midplane, the low-metallicity tail is slightly more prominent for the two-infall than the inside-out model, and the two-infall MDFs extend to slightly higher metallicity. In the outer Galaxy, though, the MDFs produced by the two models are nearly identical. The shift in the skewness and peak of the MDF from the inner to the outer Galaxy is unaffected by the choice of SFH.

Holding the SFH fixed, varying the DTD has a minimal effect on the MDFs. The two right-hand columns of Figure 5 plot the MDFs for two multizone models, which both assume an inside-out SFH but different DTDs: a power-law with slope $\alpha = -1.4$, and an exponential with timescale $\tau = 3$ Gyr. The balance between prompt and delayed SNe Ia is starkly different between the two models, with ∼80% of explosions occurring within 1 Gyr in the former but only ∼30% in the latter. However, the effect on the MDF is interestingly small, given this difference. The steep power law leads to an MDF at small $R_{\rm gal}$ that is only slightly narrower than the extended exponential (made apparent by the higher peak of the normalized MDF). This tracks with our findings from one-zone models in Section 3 that the DTD has a smaller effect on the MDF than other observables.

The inner Galaxy MDF is more sensitive to the choice of DTD than the outer Galaxy. Here, the SFH peaks earlier and declines more sharply due to the inside-out formation of the disk. Consequently, SNe Ia often explode when the gas supply is significantly lower than when the progenitors formed. This so-called "gas-starved ISM" effect drives a faster increase in metallicity (see analytic demonstration in Weinberg et al. 2017), which ultimately lowers the number of low-metallicity stars. The more extended the DTD, the stronger the effect. The outer disk is less affected by the choice of DTD, though, due to the more extended SFH.

To quantify the agreement between the MDFs generated by `VICE` and those observed in APOGEE, we compute the Kullback–Leibler (KL) divergence (Kullback & Leibler 1951), defined as

$$D_{\rm KL}(P||Q) \equiv \int_{-\infty}^{\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right) dx, \tag{18}$$

for distributions P and Q with probability density functions (PDFs) $p(x)$ and $q(x)$. If $D_{\rm KL} = 0$, the two distributions contain equal information. In this case, P is the APOGEE MDF, Q is the model MDF, and $x = [{\rm Fe/H}]$. We forward-model the observational uncertainties given in Table 5 by applying a random Gaussian scatter to the abundance of each model stellar population, and we numerically evaluate Equation (18) with integration step size $d[{\rm Fe/H}] = 0.01$ dex. For each SFH and DTD, we compute $D_{\rm KL}$ in the 18 different Galactic regions shown in Figure 5. We use bins in $R_{\rm gal}$ with a width of 2 kpc between 3 and 15 kpc, and bins in midplane distance of $|z| = 0$–0.5 kpc, 0.5–1 kpc, and 1–2 kpc. The score S for the entire model is taken to be the average of $D_{\rm KL}$ for each region

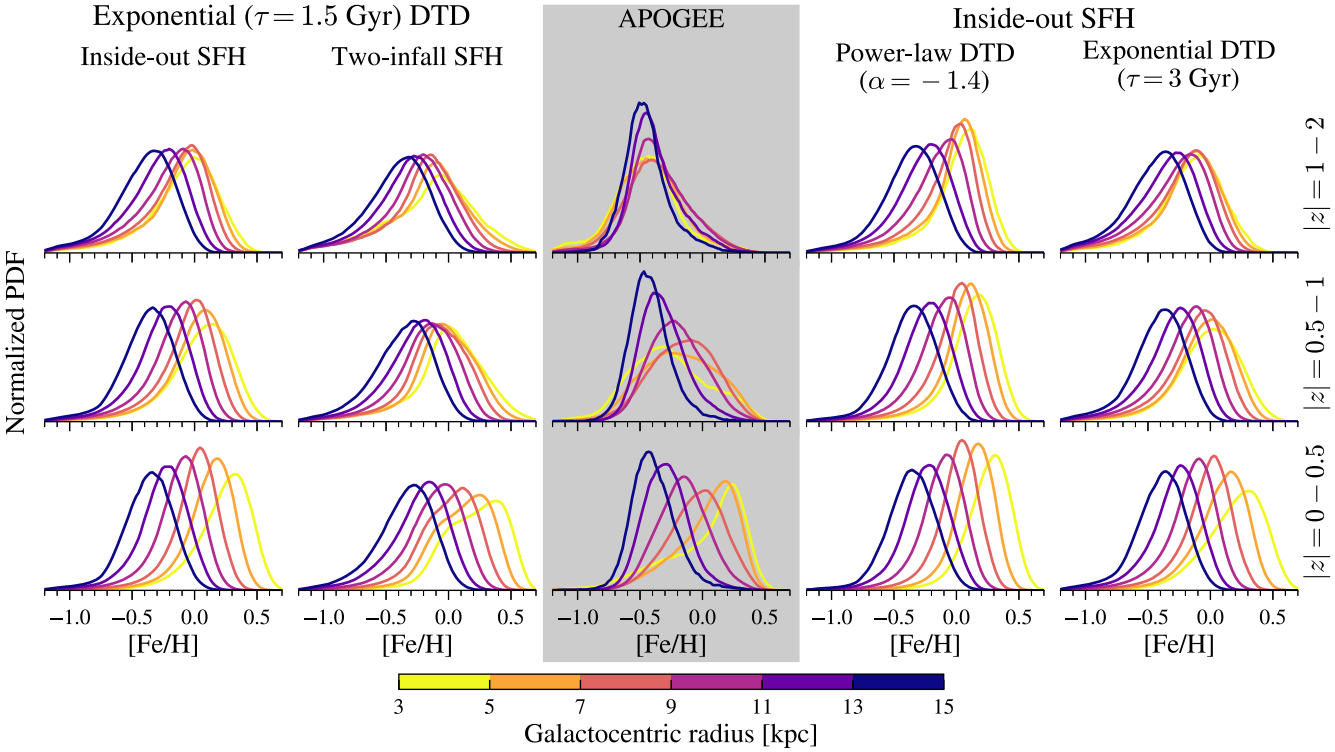

Figure 5. MDFs from multizone models with various SFHs and DTDs. Each row presents distributions of stars within a range of midplane distance: $1 \leqslant |z| < 2$ kpc (top), $0.5 \leqslant |z| < 1$ kpc (middle), and $0 \leqslant |z| < 0.5$ kpc (bottom). Within each panel, curves of different color represent the distributions of stars binned by Galactocentric radius $R_{\rm gal}$, from $3 \leqslant R_{\rm gal} < 5$ kpc (yellow) to $13 \leqslant R_{\rm gal} < 15$ kpc (blue). Each distribution is normalized so the area under the curve is 1, and the vertical scale is consistent across each row. A Gaussian scatter with a width equal to the median observational uncertainty in APOGEE DR17 (see Table 5) is applied to the abundance of each model stellar population. For visual clarity, each MDF is smoothed with a boxcar width of 0.2 dex. Left columns: comparison between the inside-out and two-infall SFHs; both assume the exponential ($\tau = 1.5$ Gyr) DTD. Center column: the distributions from APOGEE DR17 for reference, binned and smoothed similarly. Right columns: comparison between the power-law ($\alpha = -1.4$) and exponential ($\tau = 3$ Gyr) DTDs with the inside-out SFH. The MDFs in the inner Galaxy show the greatest change between the DTDs (see discussion in Section 4.1).

($R_{\rm gal}$, $|z|$) weighted by the number of APOGEE stars in that region $N_\star(R_{\rm gal}, |z|)$ (see Table 4):

$$S = \frac{\sum_{R_{\rm gal},|z|} D_{\rm KL}(P||Q|R_{\rm gal}, |z|) N_\star(R_{\rm gal}, |z|)}{\sum_{R_{\rm gal},|z|} N_\star(R_{\rm gal}, |z|)}. \quad (19)$$

The model combination with the best (lowest) score for the MDF is the two-infall SFH with the triple-system DTD. The choice of SFH has a larger effect on the overall score than the DTD, and the best-performing SFH is the two-infall model. However, the difference between the best-scoring model and the worst (inside-out SFH with the $\alpha = -1.4$ power-law DTD) is fairly small. While there are some *quantitative* differences in how the shape of the MDF varies with Galactic region, the *qualitative* trends are unaffected by the choice of model SFH or DTD. These trends are primarily driven by the assumption of chemical equilibrium, the abundance gradient, and radial migration (see discussion in Section 3.2 of J21).

4.2. The Distribution of [O/Fe]

The distribution of [O/Fe] serves as a record of the relative rates of SNe Ia and CCSNe. As such, its shape is affected by both the SFH and DTD. Figure 6 shows the distribution of [O/Fe] across the disk for the four model SFHs compared to the distributions measured by APOGEE. All four models assume an exponential DTD with $\tau = 1.5$ Gyr, which has an intermediate median delay time among all our DTDs. We see similar trends with Galactic region across all four models. Near the midplane, the distributions depend minimally on radius, but away from the midplane, there is a clear trend toward higher [O/Fe] at small $R_{\rm gal}$.

While trends with $R_{\rm gal}$ and $|z|$ are similar across the different models, the shape of the distribution varies greatly with the chosen SFH. The inside-out and late-burst models produce similar distributions because of the similarity of their underlying SFHs, as the burst is imposed upon the inside-out SFH (see Equation (12)). Both skew heavily toward near-solar [O/Fe], although the late-burst model produces a slightly broader peak and a less-prominent high-[O/Fe] tail. This difference arises because the late-burst SFH shifts a portion of the stellar mass budget to late times when [O/Fe] is low. The only region that shows any significant skew toward high [O/Fe] is $R_{\rm gal} = 3$–5 kpc and $|z| = 1$–2 kpc, but the shift to higher [O/Fe] at high latitudes is gradual and does not produce the notable trough at [O/Fe] ≈ 0.2 which is seen in the APOGEE data.

On the other hand, the early-burst model produces a bimodal [O/Fe] distribution in most regions. Although agreement is not perfect, the early-burst SFH produces the closest match to the data by far. In particular, the low-α sequence away from the midplane is dominated by stars in the solar annulus and outer disk, a trend that is also seen in APOGEE. However, the early-burst high-α sequence contains many stars in the outer disk and close to the midplane, whereas the APOGEE distribution does not show a prominent high-α peak beyond $R_{\rm gal} \sim 11$ kpc ($\sim$7 kpc in the midplane).

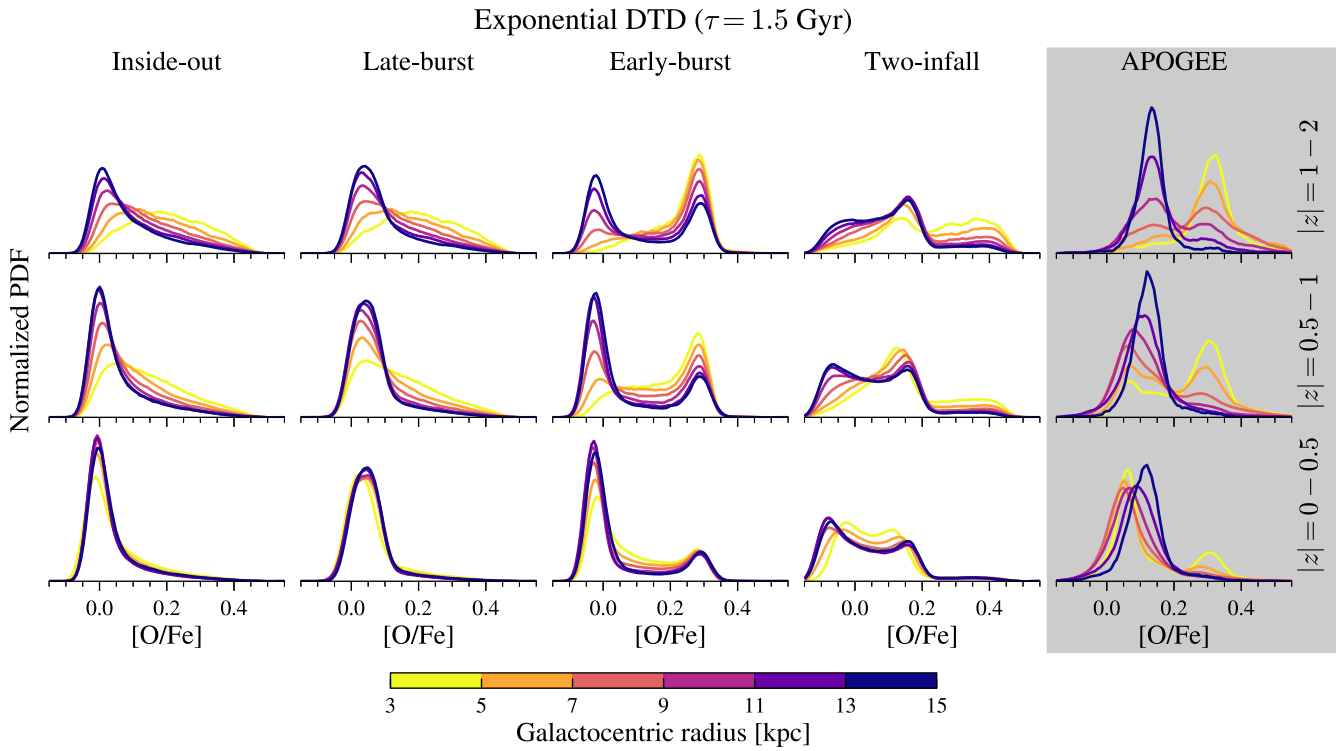

Figure 6. Distributions of [O/Fe] from multizone models with different SFHs. All assume the exponential ($\tau = 1.5$ Gyr) DTD. The format of each panel is the same as in Figure 5, except that all distributions are smoothed with a boxcar width of 0.05 dex. Distributions from APOGEE DR17, binned and smoothed similarly, are presented in the rightmost column for reference.

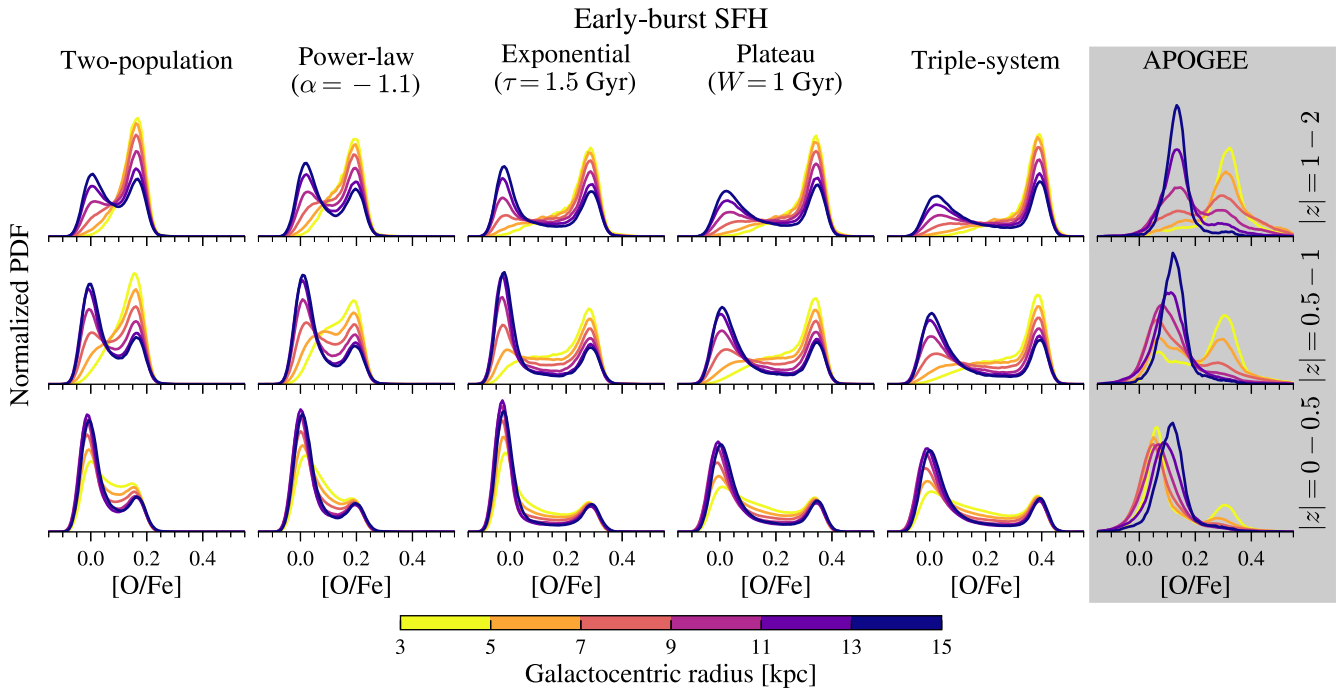

Figure 7. The same as Figure 6 but for different DTDs. In all cases, an early-burst SFH is assumed.

The two-infall SFH produces *three* distinct modes at [O/Fe] ≈ -0.05, 0.15, and 0.4. At small $R_{\rm gal}$ and with increasing $|z|$, the low-α peak decreases in prominence as the high-α peak increases, but the intermediate peak is a striking feature at all latitudes that does not align with observations. In the APOGEE data, the high-α peak is at [O/Fe] ≈ 0.3, roughly halfway between the intermediate and high-α peaks produced by the two-infall model. However, the model high-α sequence does match the observed trends with $R_{\rm gal}$ and $|z|$ much better than the early-burst models.

Figure 7 shows [O/Fe] distributions produced by models with the same SFH but a range of different DTDs. We show models with the early-burst SFH because it produces distinct low- and high-α sequences. The most obvious effect of the

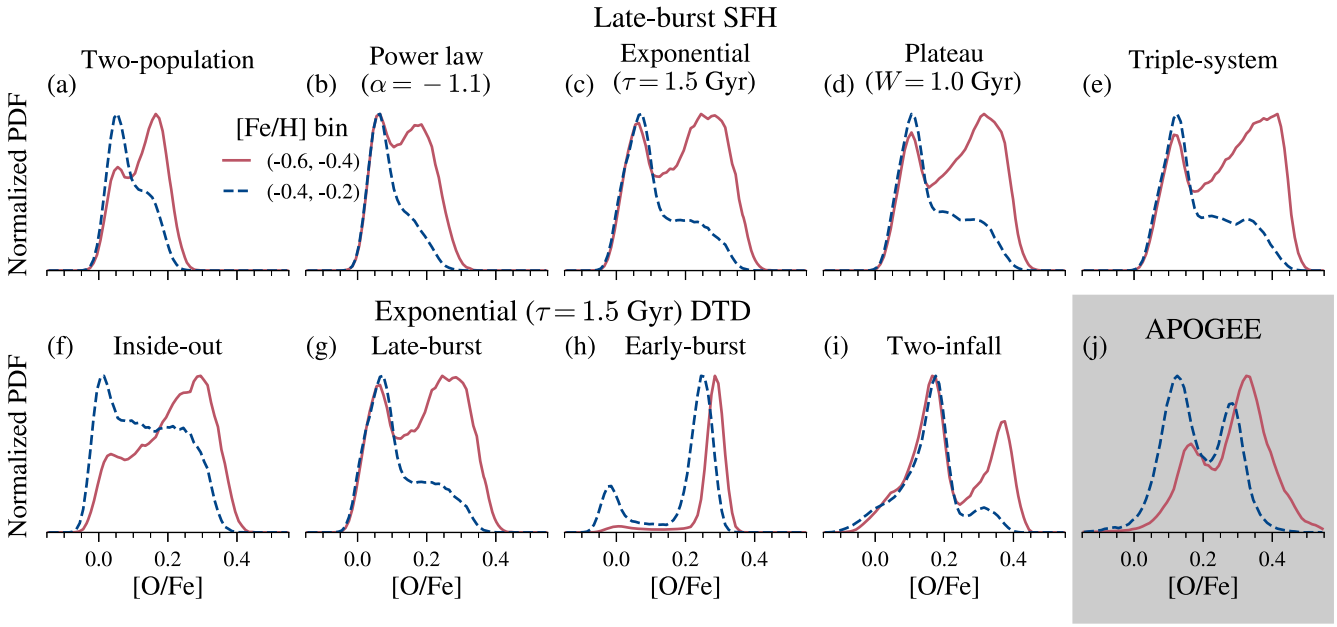

Figure 8. The distributions of [O/Fe] along two different slices of [Fe/H]: $-0.6 \leqslant$ [Fe/H]< -0.4 (red solid) and $-0.4 \leqslant$ [Fe/H]< -0.2 (blue dashed). Each panel contains stars within the Galactic region defined by $7 \leqslant R_{\rm gal} < 9$ kpc and $0 \leqslant |z| < 2$ kpc. For each distribution, 100,000 stellar populations are resampled from the model output to match the $|z|$ distribution of the APOGEE sample. Top row: results from five multizone models that assume the late-bust SFH but different DTDs. Bottom row: the first four panels compare the four SFHs (see Figure 2), all assuming an exponential DTD with $\tau = 1.5$ Gyr. The bottom-right panel (highlighted) plots data from APOGEE DR17 for reference.

DTD is to shift the mode of the high-α sequence. The two-population DTD, which has the most prompt SNe Ia, places the high-α sequence at [O/Fe] ≈ 0.15, while the triple-system DTD, which has the fewest prompt SNe Ia, places it $\sim$0.25 dex higher at [O/Fe] ≈ 0.4. The plateau ($W = 1$ Gyr) DTD places the higher peak at [O/Fe] ≈ 0.35, close to where it appears in the APOGEE distributions. However, the distance *between* the peaks of the APOGEE distributions is only $\sim$0.2 dex, since the observed low-α sequence sits at [O/Fe] ≈ 0.1. This spacing is best replicated by the power-law ($\alpha = -1.1$) DTD, even though both peaks sit $\sim$0.1 dex too low and the distributions are narrower than observed.

In general, models with fewer prompt SNe Ia populate the high-α sequence with more stars because the chemical evolution track spends more time in the high-α regime. This qualitatively agrees with the isolated and cosmological simulations of Poulhazan et al. (2018), who find that DTDs with a significant prompt component produce narrower [O/Fe] distributions and a higher average [O/Fe].

We again compute the KL divergence (Equation (18)) to quantify the agreement between the [O/Fe] DFs of our models and APOGEE. We calculate a score for each model as described in Section 4.1. The best-scoring model combines the inside-out SFH with the triple-system DTD, and the plateau ($W = 1$ Gyr) and exponential ($\tau = 3$ Gyr) DTDs score well when combined with either the inside-out or late-burst SFHs. Both plateau DTDs also score relatively well with the two-infall SFH. Surprisingly, the early-burst SFH scores quite poorly for all DTD models, despite the fact that it produces the most distinct high- and low-α sequences. We discuss this further in Section 5.1.

4.3. Bimodality in [O/Fe]

The [O/Fe] distributions from APOGEE in Section 4.2 show two distinct peaks whose relative prominence varies with $R_{\rm gal}$ and $|z|$ (see also Figure 4 of Hayden et al. 2015). A crucial feature of this bimodality, which is not apparent in the analysis of the previous section, is the presence of both sequences at fixed [Fe/H]. The separation between the two sequences appears to be a real feature and not an artifact of the APOGEE selection function (Vincenzo et al. 2021). A successful model for the evolution of the Milky Way therefore must reproduce this bimodality.

Figure 8 compares the [O/Fe] distributions in the solar annulus ($7 \leqslant R_{\rm gal} < 9$ kpc and $0 \leqslant |z| < 2$ kpc) in two bins of [Fe/H] ($-0.6 <$ [Fe/H] < -0.4 and $-0.4 <$ [Fe/H] < -0.2) for select model outputs and APOGEE data. The purpose of the narrow [Fe/H] bins is to isolate the bimodality of the [O/Fe] distribution with minimal variation in [Fe/H]. The APOGEE distributions in the bottom-right panel (j) show that the high-α mode is more prominent at lower [Fe/H], but the distributions in both bins are clearly bimodal. The "trough" occurs near [O/Fe] ≈ 0.2 in each bin.

To quantify the strength of the α-bimodality, we use the peak-finding algorithm `scipy.signal.find_peaks` (Virtanen et al. 2020). For each peak, we calculate the prominence, or the vertical distance between a peak and its highest neighboring trough. We consider a distribution bimodal if both peaks exceed an arbitrary threshold of 0.1. The APOGEE distributions exceed this threshold in both [Fe/H] bins.

The top row of Figures 8(a)–(e) shows the [O/Fe] bimodality (or lack thereof) across five different DTDs, all of which assume the late-burst SFH. To better approximate the APOGEE selection function, we resample our model stellar populations so the $|z|$ distribution closely matches that of APOGEE in the solar neighborhood. Six of the eight DTDs (all except the two-infall and $\alpha = -1.4$ power-law DTDs) exceed our prominence threshold in the low-[Fe/H] bin. Panel (a) shows that the two-infall DTD produces a marginal low-α

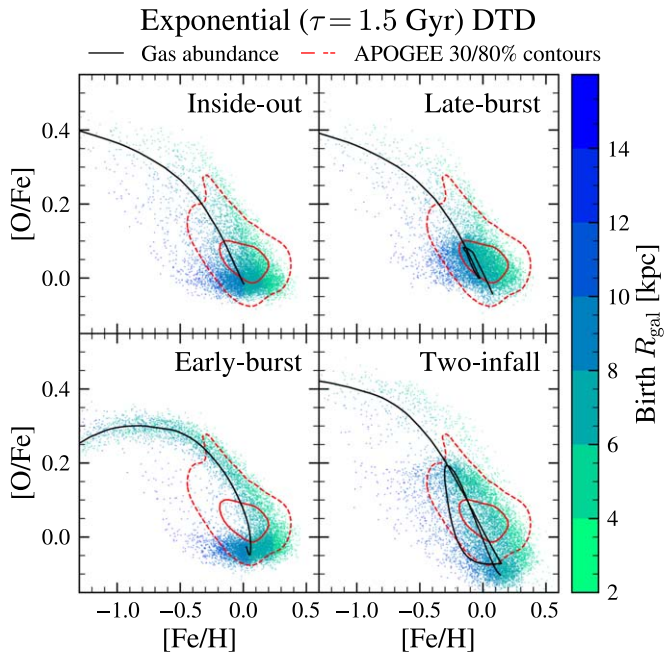

Figure 9. A comparison of the [O/Fe]–[Fe/H] plane between the four SFHs in our multizone models. All assume the exponential ($\tau = 1.5$ yr) DTD. Each panel plots a random mass-weighted sample of 10,000 star particles in the solar neighborhood ($7 \leqslant R_{gal} < 9$ kpc, $0 \leqslant |z| < 0.5$ kpc) color coded by R_{gal} at birth. A Gaussian scatter has been applied to all points based on the median abundance errors in APOGEE DR17 (see Table 3). The black curves represent the ISM abundance tracks in the 8 kpc zone. The red contours represent a 2D Gaussian kernel density estimate of the APOGEE abundance distribution in that Galactic region with a bandwidth of 0.03. The solid and dashed contours enclose 30% and 80% of stars in the sample, respectively.

peak, although it does not meet the prominence threshold. In general, DTDs with fewer prompt SNe Ia produce a high-α peak that is more prominent and at a higher [O/Fe], as was the case with the [O/Fe] distributions in Section 4.2.

Figures 8(f)–(i) illustrate the effect of the SFH on the [O/Fe] bimodality. The inside-out SFH does not produce a bimodal distribution for most of our DTDs (the exception is the $W = 1$ Gyr plateau DTD, which produces a much smaller trough than observed). On the other hand, the early-burst SFH *always* produces a bimodal distribution in the high-[Fe/H] bin regardless of the assumed DTD, but not in the low-[Fe/H] bin (the small low-α peak falls below our prominence threshold). For models with the late-burst and two-infall SFHs, the bimodality in the low-[Fe/H] bin is variable depending on the DTD: those with longer median delay times (e.g., exponential, plateau, or triple-system) generally produce a bimodal distribution, while the two DTDs with the most prompt SNe Ia do not.

One major problem in all of our models is the presence of the [α/Fe] bimodality across only a narrow range of [Fe/H]. Even our most successful models can produce a bimodal [O/Fe] distribution in only one bin: the high-[Fe/H] bin for the early-burst SFH, and the low-[Fe/H] bin for the late-burst and two-infall SFHs. In APOGEE, the two sequences are co-extant between [Fe/H] ≈ -0.6, below which the high-α sequence dominates, and [Fe/H] $\approx +0.2$, at which point they join. The failure of these models to fully reproduce the bimodality across the whole range of [Fe/H] was noted by J21, and the problem persists for each model we consider here.

4.4. The [O/Fe]–[Fe/H] Plane

In Section 3, we illustrated that the form and parameters of the DTD have an important effect on the ISM abundance tracks in idealized one-zone models (see Figures 3 and 4). However, comparisons to data are limited because the tracks neither record the number of stars that formed at each abundance nor incorporate the effect of stellar migration. Here, we present the distribution of stellar abundances in the [O/Fe]–[Fe/H] plane alongside the ISM abundance tracks from our multizone models. We compare our model outputs to the observed distributions from APOGEE across the Milky Way disk.

Figure 9 compares the [O/Fe]–[Fe/H] plane in the solar neighborhood ($7 \leqslant R_{gal} < 9$ kpc, $0 \leqslant |z| < 0.5$ kpc) between our four model SFHs. The black curves represent the ISM abundance as a function of time in the $R_{gal} = 8.0$–8.1 kpc zone; in the absence of radial migration, all model stellar populations would lie close to these lines. Stellar populations to the left of the abundance tracks were born in the outer disk, while those to the right were born in the inner disk, as illustrated by the color coding in the figure. Much of the scatter in [Fe/H] in a given Galactic region can be attributed to radial migration (Edvardsson et al. 1993).

The tracks predicted by all four SFHs initially follow a similar path of decreasing [O/Fe] with increasing [Fe/H]. The ISM abundance ratios of the inside-out model change monotonically over the entire disk lifetime. The stellar abundance distribution at both low- and high-[O/Fe] is composed of stars with a wide range of birth R_{gal}.

The late-burst model produces similar results to the inside-out model up to [Fe/H] ≈ -0.2, due to their similar SFHs. The Gaussian burst in its SFH introduces a loop in the ISM abundance track, as an uptick in star formation at $t \approx 11$ Gyr raises the CCSN rate, leading to a slight increase in [O/Fe] before the subsequent increase in the SN Ia rate lowers the [O/Fe] once again (see, e.g., Figure 1 of Johnson & Weinberg 2020). This loop slightly broadens the low-[O/Fe] stellar distribution, as we observed in Section 4.2.

This same pattern is seen much more strongly in the abundance tracks for the two-infall model. Here, the significant infall of pristine gas at $t = 4$ Gyr leads to rapid dilution of the metallicity of the ISM, followed by a large burst in the SFR, which raises [O/Fe] by $\sim$0.2 dex. We observe a ridge in the stellar abundance distribution at the turn-over point ([O/Fe] ≈ 0.15) associated with SNe Ia whose progenitors formed during the burst. This ridge roughly coincides with the upper limit of the APOGEE distribution near the midplane. The three-peaked structure of the [O/Fe] distributions in Section 4.2 is explained by the abundance tracks here: a small population of stellar populations at [O/Fe] ≈ 0.4 is produced initially, followed by the middle peak when the abundance track turns over, and finally the peak at [O/Fe] ≈ -0.1, which reflects the equilibrium abundance ratio of the second infall.

The early-burst track is the most distinct from the other models at low metallicity. The portion shown in Figure 9 represents the evolution *after* the early SFE burst. At low metallicity, there is a "simmering phase" where [O/Fe] slowly decreases to a local minimum at [Fe/H] ≈ -1.3, at which point the rapid increase in the SFE causes the [O/Fe] to rebound (a more thorough examination of this behavior can be found in Conroy et al. 2022). The early-burst SFH produces the clearest separation between a high- and low-[O/Fe] sequences. The number of stars on the high-[O/Fe] sequence is relatively high,

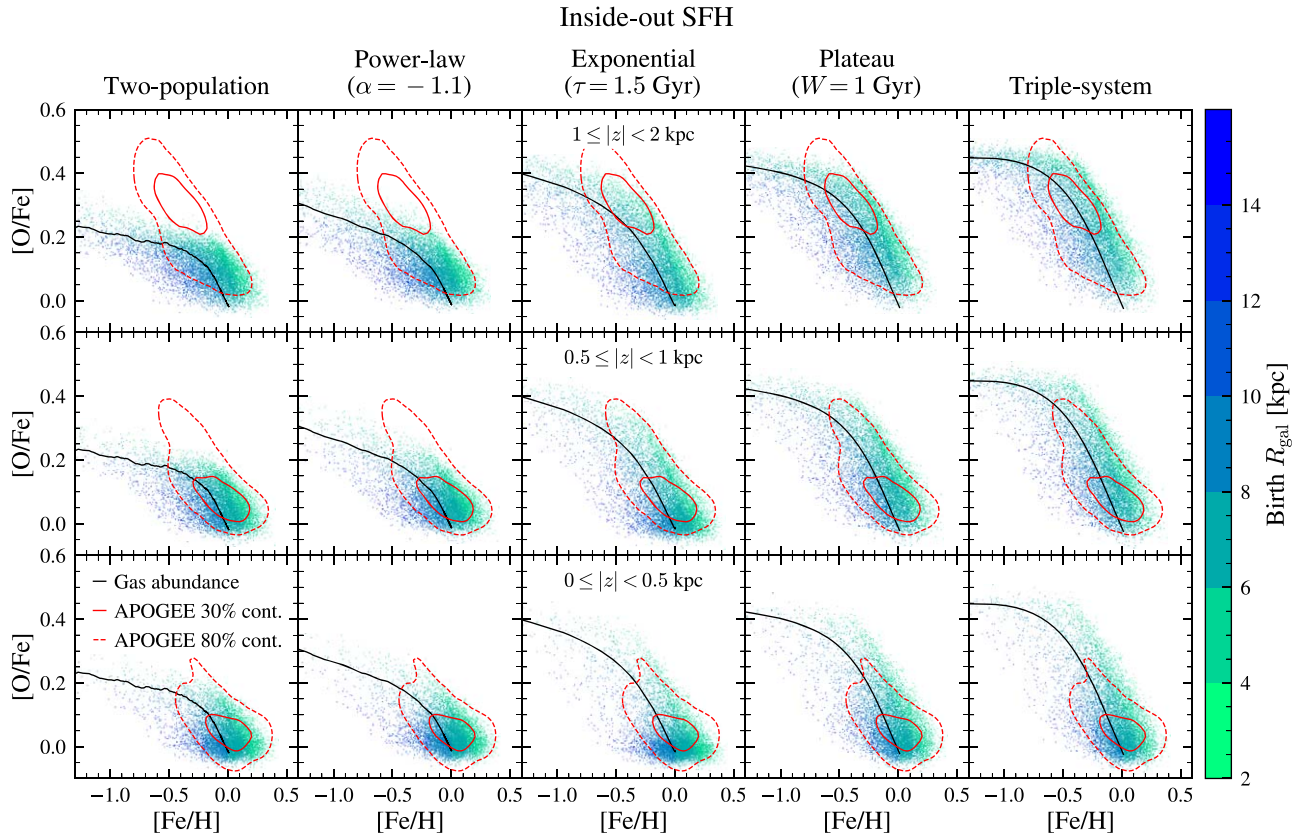

Figure 10. The [O/Fe]–[Fe/H] plane from multizone models with different DTDs (see Figure 1). All assume the inside-out SFH. Each panel is similar to those in Figure 9, except each row contains star particles from a different bin in $|z|$, with stars closest to the midplane in the bottom row and stars farthest from the midplane in the top row as labeled in the middle column. All panels contain stars within the solar annulus ($7 \leqslant R_{gal} < 9$ kpc).

likely as a result of its higher SFR at early times compared to the other models.

Figure 10 compares the [O/Fe]–[Fe/H] ISM tracks and stellar distributions for five models with the same SFH but different DTDs. We choose the inside-out SFH for this figure because it predicts monotonically decreasing abundance ratios, making comparisons between the different DTDs relatively straightforward. The models are arranged according to the median delay time of the DTD, increasing across the panel columns from left to right.

The two-population and power-law ($\alpha = -1.1$) DTDs, which have a large fraction of prompt ($t \lesssim 100$ Myr) SNe Ia, produce stellar abundance distributions that are reasonably well aligned with the APOGEE contours at low $|z|$, but they entirely miss the observed high-α sequence at large $|z|$. The ISM abundance tracks for the 8 kpc zone do not pass through the APOGEE 30% contour at $|z| = 1$–2 kpc. For both DTDs, the high-[O/Fe] knee is located below the leftmost bound of the plot, but we observe a second knee at [O/Fe] ≈ 0.15 where the abundance tracks turn downward once more. As discussed in Section 3, the second knee is most prominent in the model with the two-population DTD because of its long exponential tail.

The exponential ($\tau = 1.5$ Gyr) DTD, which has an intermediate median delay time, produces a distribution in Figure 10 that aligns quite well with the 80% APOGEE contours in all $|z|$-bins, and even produces a "ridge" that extends to high [O/Fe] at low- and mid-latitudes (bottom and center panels, respectively). While it does better at populating the high-α sequence than the previous DTDs, the bulk of the model stellar populations at large $|z|$ still fall below the APOGEE 30% contour.

The two right-hand columns present model results for the plateau ($W = 1$ Gyr) and triple-system DTDs, which have the longest median delay times. The high-[O/Fe] knee occurs at a much higher metallicity in these models and is visible in the gas abundance tracks in the upper-left corner of the panels. At large $|z|$, the predicted abundance distributions align quite well with the APOGEE high-α sequence, but there is a significant ridge of high-α stars from the inner Galaxy at low $|z|$.

To quantify the agreement between the multizone model outputs and data in [O/Fe]–[Fe/H] space, we implement the method of Perez-Cruz (2008) for estimating the KL divergence between two continuous, multivariate samples using a k-nearest neighbor estimate. For n samples from a multivariate PDF $p(\boldsymbol{x})$ and m samples from $q(\boldsymbol{x})$, we can estimate $D_{\mathrm{KL}}(P||Q)$ according to the following:

$$\hat{D}_k(P||Q) = \frac{d}{n}\sum_{i=1}^{n} \log \frac{r_k(\boldsymbol{x}_i)}{s_k(\boldsymbol{x}_i)} + \log \frac{m}{n-1}, \tag{20}$$

where $d = 2$ is the dimension of the sample space and $r_k(\boldsymbol{x}_i)$ and $s_k(\boldsymbol{x}_i)$ are the distance to the kth nearest neighbor of $\boldsymbol{x}_i$ in the samples of P and Q, respectively. We take $k = 2$ to find the nearest neighbor other than the sample itself. As before, P is the APOGEE distribution and Q is the model distribution, and in this case $\boldsymbol{x} = ([\mathrm{Fe/H}], [\mathrm{O/Fe}])$, without applying any scaling factor to each dimension. As in Sections 4.1 and 4.2, we bin the model outputs and data by R_{gal} and $|z|$, calculate $\hat{D}_k(P||Q)$ in

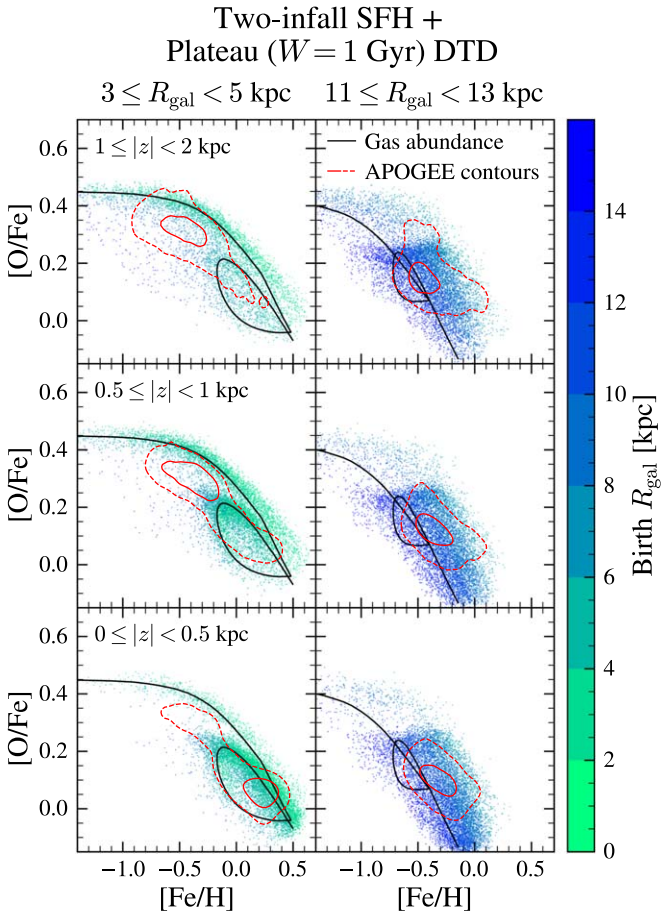

Figure 11. The [O/Fe]–[Fe/H] plane for multiple Galactocentric regions from the model with the two-infall SFH and plateau ($W = 1$ Gyr) DTD. The two columns of panels contain stars in different bins of $R_{\rm gal}$, and each row contains stars from a different bin of $|z|$. The contents of each panel are as described in Figure 9.

each region, and then take the weighted mean of each region as in Equation (19) to arrive at a single score for each model.

The best-scoring model combines the triple-system DTD with the inside-out SFH. The two other DTDs with the longest median delay times, plateau ($W = 1$ Gyr) and exponential ($\tau = 3$ Gyr), also score quite well. As with the [O/Fe] DFs, the inside-out and late-burst SFH models score similarly across all DTDs in the [O/Fe]–[Fe/H] plane. The early-burst SFH models score the worst out of all the SFHs, likely due to the long "tail" in the distribution down to low [Fe/H] which is not seen in the APOGEE data.

Figure 11 plots the stellar [O/Fe]–[Fe/H] abundances from the model with the two-infall SFH and plateau ($W = 1$ Gyr) DTD in two different bins of $R_{\rm gal}$. In the inner Galaxy, the model distribution at large $|z|$ lies at higher [O/Fe] and is more extended than the APOGEE distribution. Agreement between the model and data is worst at mid-latitudes: the model distribution is sparsest in the area of the peak of the APOGEE distribution. Near the midplane, however, the model output is well aligned with the data. In the outer Galaxy, the distributions are well aligned at all $|z|$, though the model distributions are more extended along the [O/Fe] axis than in the data. Adjustments to the yields or the relative infall strengths could improve the agreement between the two-infall model output and the observed distributions.

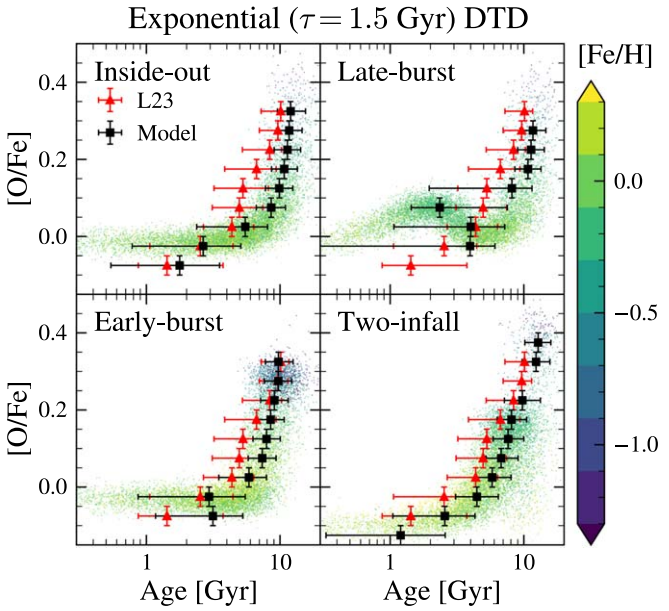

Figure 12. A comparison of the age–[O/Fe] relation between multizone models with different SFHs. All assume the exponential ($\tau = 1.5$ Gyr) DTD. Each panel plots a random mass-weighted sample of 10,000 star particles in the solar neighborhood ($7 \leqslant R_{\rm gal} < 9$ kpc, $0 \leqslant |z| < 0.5$ kpc) color coded by [Fe/H]. A Gaussian scatter has been applied to all points based on the median [O/Fe] error from APOGEE DR17 and the median age error from L23 (see Table 3). Black squares represent the mass-weighted median age of star particles within bins of [O/Fe] with a width of 0.05 dex, and the horizontal black error bars encompass the 16th and 84th percentiles. Red triangles and horizontal error bars represent the median, 16th, and 84th percentiles of age from L23, respectively. For clarity, bins that contain less than 1% of the total mass (in the models) or total number of stars (in the data) are not plotted.

4.5. The Age–[O/Fe] Plane

As demonstrated by our one-zone models in Section 3, models that produce similar tracks in abundance space can be distinguished by the rate of their abundance evolution. We therefore expect the age–[O/Fe] relation to be a useful diagnostic. Figure 12 shows the stellar age and [O/Fe] distributions in the solar neighborhood for each of our four SFHs. As in Figure 9, all four panels assume an exponential DTD with $\tau = 1.5$ Gyr. We compare these predictions against ages estimated with L23's variational encoder–decoder algorithm. We caution against drawing strong conclusions from this comparison, because we do not correct for selection effects or systematic errors in the age determination.

The inside-out and late-burst models show fair agreement with the data at high [O/Fe], although both show a $\sim$2 Gyr offset. One could shift the ramp-up in star formation to slightly later times or simply run the model for a shorter amount of time to close this gap. Although it is a visually striking difference, the age at the high-[O/Fe] knee is not a good diagnostic for the SFH after factoring in the age uncertainties. As in the data, the trend in the median age with decreasing [O/Fe] decreases monotonically in the inside-out model. The late-burst model, however, shows a bump in the relation at a look-back time of $\sim$2 Gyr, which is not seen in the data, as noted by J21.

For the early-burst SFH, the predicted stellar ages are almost perfectly aligned with the data for [O/Fe] $\gtrsim 0.2$. The rapid rise in the SFE at early times delays the descent to lower [O/Fe] values and produces a clump of low-metallicity, high-[O/Fe] stars at an age of $\sim$10 Gyr. Finally, the two-infall SFH

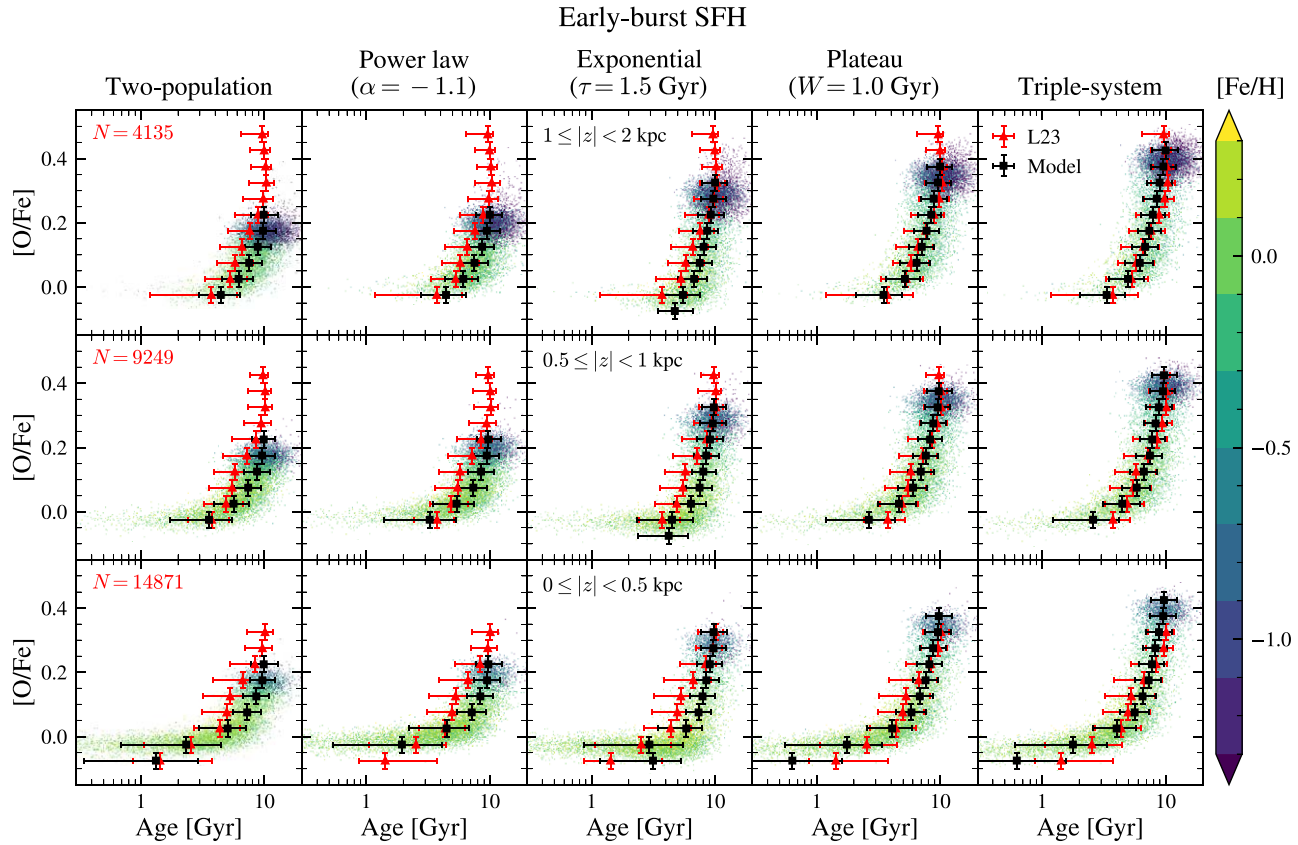

Figure 13. A comparison of the age–[O/Fe] relation between multizone models with different DTDs. All assume the early-burst SFH. Each row contains star particles from a different bin in $|z|$, with stars closest to the midplane in the bottom row and stars farthest from the midplane in the top row as labeled in the middle column. In all panels stars are limited to the solar annulus ($7 \leqslant R_{\rm gal} < 9$ kpc), and the layout of each panel is the same as in Figure 12. The red numbers in the leftmost panels indicate the number of APOGEE stars with L23 ages in each bin of $|z|$.

produces a fair match to the data. Stars with [O/Fe] $\gtrsim 0.25$ were produced in the first infall, while the second infall produces a clump of stars with similar metallicity, ages of $\sim$8 Gyr, and [O/Fe] ≈ 0.2. There is a population of old, low-α stars that arise due to the initial descent in [O/Fe] prior to the second accretion epoch. The subsequent increase in [O/Fe] does not produce as strong of a bump as the late-burst SFH, because it occurs much earlier and is therefore narrower in log (age). However, the two-infall SFH produces [O/Fe] abundances for the youngest stars that are roughly 0.1 dex lower than the other models.

In contrast to J21, none of our models predict a population of young, α-enhanced stars in the solar neighborhood. These stars have been observed in APOGEE (e.g., Martig et al. 2016; Silva Aguirre et al. 2018) and many are likely old systems masquerading as young stars due to mass transfer or a merger (e.g., Yong et al. 2016), but it is not known whether some fraction are truly intrinsically young (Hekker & Johnson 2019). In J21, these young, α-enhanced stars are the result of a highly variable SN Ia rate in the outer Galaxy. The SN Ia progenitors migrate before they are able to enrich their birth annulus, so the subsequent stellar populations are depleted in Fe. Two differences in the migration scheme explain the lack of these stars in our own models: first, we adopt a time dependence for radial migration of $\Delta t^{1/3}$, which is slower than the diffusion scheme ($\Delta t^{1/2}$) of J21. Second, our migration method is designed to produce smooth abundance distributions, whereas the method of J21 can assign identical migration patterns to many stellar populations in sparsely populated regions of the Galaxy, potentially removing many SN Ia progenitors from a given zone simultaneously (for more discussion, see Appendix C). This update to the model is consistent with Grisoni et al.'s (2024) finding that young α-rich stars have similar occurrence rates across the disk, which supports a stellar, as opposed to Galactic, origin.

Figure 13 shows the predicted age–[O/Fe] relation for five of our DTDs. All models were run with the early-burst SFH because it predicts the clearest separation between the high- and low-α sequences (see Figure 12). Similar to Figure 10, models are arranged from left to right by increasing median SN Ia delay time. The high-α sequence moves to higher [O/Fe] with increasing median delay time, from $\sim$0.2 for the two-population model to $\sim$0.4 for the triple-system DTD. As we have seen in previous figures, the range in [O/Fe] produced by DTD models with many prompt SNe Ia is much smaller than the extended DTDs. At high $|z|$ (top row), the observed range of [O/Fe] is larger than what is produced by most of our models. While the plateau ($W = 1$ Gyr) and triple-system models come close, the other three fall short of the observed range in [O/Fe], but still closely match the median age–[O/Fe] relation. There is a slight reversal in the observed trend for the stars with the highest [O/Fe]: the $0.45 \leqslant$ [O/Fe] < 0.5 bin has a slightly *lower* median age than the $0.3 \leqslant$ [O/Fe] < 0.35 bin at high $|z|$ in the L23 sample, a small effect but one that is not predicted by any of our models.

Moving to stars at low $|z|$, the plateau ($W = 1$ Gyr) and triple-system DTDs overproduce stars at the old, high-α end of the distribution, while also diverging somewhat from the

observed sequence near solar [O/Fe]. The exponential ($\tau = 1.5$ Gyr) DTD comes closest to reproducing the observed range in [O/Fe], while the two DTDs with the shortest median delay time once again produce a smaller range of [O/Fe] than observed. We note that the break between the linear and flat parts of the relation is sharpest for the exponential DTD, and a more gradual transition is observed for the other four DTDs. This difference arises because the exponential DTD is most dominant at intermediate delay times ($t \sim 1$–3 Gyr) but falls off much faster than the other models at long delay times, so [O/Fe] is close to constant for look-back times $\lesssim$5 Gyr. Overall, the exponential ($\tau = 1.5$ Gyr) DTD most closely matches the data for stars with $0 \leqslant |z| < 0.5$ kpc.

Here, we use a different scoring system than in previous subsections, due to the much larger uncertainties in age than [O/Fe]. As shown in Figures 12 and 13, in each Galactic region, we sort the model outputs and data into bins of [O/Fe] with a width of 0.05 dex. We define the rms median age difference for the region as

$$\Delta\tau_{\rm RMS} \equiv \sqrt{\frac{\sum_k \Delta\tau_k^2 n_{{\rm L23},k}}{n_{\rm L23,tot}}}, \tag{21}$$

where $\Delta\tau_k = {\rm med}(\tau_{\rm VICE}) - {\rm med}(\tau_{\rm L23})$ is the difference between the mass-weighted median age in VICE and the median stellar age from L23 in bin k, $n_{{\rm L23},k}$ is the number of stars from the L23 age sample in bin k, and $n_{\rm L23,tot}$ is the total number of stars in the sample in that Galactic region. This is similar to a reduced χ^2 estimator, except that the difference in medians is not weighted by the variance in the observed sample. If a bin has no modeled or observed stars, we do not calculate $\Delta\tau_k$ for that bin. As before, the score for the model as a whole is the average of $\Delta\tau_{\rm RMS}$ across all regions, weighted by the number of stars with age measurements in each region.

The best (lowest-scoring) model in the age–[O/Fe] plane is the triple-system DTD with the two-infall SFH. Models that score almost as well are the plateau ($W = 1$ Gyr) and triple-system DTDs with either the early-burst or two-infall SFH. Visually, the non-monotonic bump in the age–[O/Fe] relation produced by the late-burst SFH does not match the observed distribution, but it actually *improves* $\Delta\tau_{\rm RMS}$ by lowering the median age of stars in the low-[O/Fe] bins. If the shape of the distribution is taken into account, the late-burst SFH produces the worst match to the data. We discuss the quantitative scores in further Section 5.1 below.

5. Discussion

5.1. Qualitative Comparisons

In Section 4, we focused on a representative subset of our 32 multizone models (four SFHs and eight DTDs). Here, we compare all of our model outputs to APOGEE across five observables: the MDF, [O/Fe] DF, [O/Fe]–[Fe/H] plane, age–[O/Fe] plane, and [O/Fe] bimodality. We perform statistical tests between APOGEE and the model outputs in each region of the Galaxy as described in corresponding subsections of Section 4, then compute the average weighted by the size of the APOGEE sample in each region to obtain a single numerical score.

The relative performance of each model is summarized in Table 6. We use these scores to indicate combinations of SFH and DTD that are favorable or unfavorable in certain regimes, but we do *not* fit our models to the data, due to computational expense. To avoid drawing strong conclusions from small numerical differences in scores, we simply write ✓, $\sim$, or $\times$, which corresponds to a score in the top, middle, or bottom third out of all models, respectively. The exact numerical scores are presented in Appendix D.

Some of the variation between models can be explained by the choice of SFH. The two-infall models tend to outperform the others for the MDFs, while the late-burst models score poorly, especially with the prompt DTDs. The early-burst models consistently have the lowest scores for the [O/Fe] DF and [O/Fe]–[Fe/H] distribution, but are able to produce a bimodal [O/Fe] distribution with every DTD (see discussion in Section 4.3). The late-burst and two-infall SFHs also produce a bimodal [O/Fe] distribution with all DTDs except those with the highest prompt fraction, while the inside-out models never produce bimodality. The inside-out models also tend to score poorly in the age–[O/Fe] plane, while the early-burst models tend to score well, although as discussed in Section 4.5, adjusting the time of the peak SFR or running the models for a shorter period of time would affect the level of agreement in the high-[O/Fe] bins.

It is somewhat surprising that the early-burst models score poorly against the APOGEE [O/Fe] DFs, given that they produce the clearest bimodal distributions. The KL divergence test heavily penalizes models with a high density in a region where the observations have little, as is the case for the high-α sequence in the outer Galaxy and close to the midplane (see Figure 6). This similarly explains the early-burst models' poor performance in the [O/Fe]–[Fe/H] plane. An iteration of this SFH where the early burst predominantly affects the inner galaxy is probably more accurate and might have more success at reproducing the [O/Fe] DF across the disk.

The choice of DTD has a clear effect on the model scores, and this effect is similar for most of the observables. The models that perform the best (most ✓ marks and fewest $\times$ symbols) are the most extended DTDs with the fewest prompt SNe Ia: both plateau DTDs, the exponential DTD with $\tau = 3$ Gyr, and the triple-system DTD. The latter actually produces the highest scores for each observable, but the plateau DTD with $W = 1$ Gyr is the most successful across all SFHs; both models have some of the longest median delay times. Models with a large fraction of prompt SNe Ia, such as the power-law and two-population DTDs, fare quite poorly, with the steepest power-law ($\alpha = -1.4$) and two-population DTDs ending up in the bottom third across the board for most of our SFHs. The fiducial power-law ($\alpha = -1.1$) does slightly better, but still compares poorly to the more extended DTDs.

Each DTD tends to score similarly across the board, but there are some combinations of SFH and DTD that buck the general trend. For example, the two-population DTD with the early-burst SFH produces an MDF that scores relatively well. The early-burst models generally produce MDFs in the $\sim$ category, so a small increase in the numerical score bumps it up to ✓; this indicates the insensitivity of the MDF to the DTD in general. The exponential DTD with $\tau = 1.5$ Gyr has generally middling performance, but does a notably poorer job when combined with the early-burst SFH, a result of the generally poor performance of that SFH.

The plateau DTD with $W = 1$ Gyr, our most successful model overall, poorly reproduces the MDF with the late-burst

Table 6
Qualitative Summary of Comparisons between the Model Output, APOGEE DR17 Abundances, and L23 Ages for Each Multizone Model

DTD	SFH	MDF	[O/Fe] DF	[O/Fe] Bimodality	[Fe/H]–[O/Fe]	Age–[O/Fe]
Two-population	Inside-out	×	×	×	×	×
($t_p = 0.05$ Gyr)	Late-burst	×	×	×	×	×
	Early-burst	✓	×	✓	×	∼
	Two-infall	∼	∼	×	×	×
Power-law	Inside-out	×	×	×	×	×
($\alpha = -1.4$)	Late-burst	×	×	×	×	×
	Early-burst	∼	×	✓	×	×
	Two-infall	×	×	×	×	×
Power-law	Inside-out	×	∼	×	∼	×
($\alpha = -1.1$)	Late-burst	×	×	✓	∼	×
	Early-burst	∼	×	✓	×	∼
	Two-infall	✓	∼	✓	∼	×
Exponential	Inside-out	∼	∼	×	∼	×
($\tau = 1.5$ Gyr)	Late-burst	∼	✓	✓	∼	∼
	Early-burst	×	×	✓	×	∼
	Two-infall	✓	∼	✓	∼	∼
Exponential	Inside-out	✓	✓	×	✓	∼
($\tau = 3.0$ Gyr)	Late-burst	∼	✓	✓	✓	✓
	Early-burst	∼	×	✓	∼	✓
	Two-infall	✓	∼	✓	∼	✓
Plateau	Inside-out	∼	✓	×	✓	∼
($W = 0.3$ Gyr)	Late-burst	×	✓	✓	✓	∼
	Early-burst	∼	∼	✓	×	✓
	Two-infall	✓	✓	✓	✓	∼
Plateau	Inside-out	✓	✓	✓	✓	∼
($W = 1.0$ Gyr)	Late-burst	×	✓	✓	✓	✓
	Early-burst	✓	∼	✓	∼	✓
	Two-infall	✓	✓	✓	✓	✓
Triple-system	Inside-out	✓	✓	×	✓	✓
($t_{\rm rise} = 0.5$ Gyr)	Late-burst	×	✓	✓	✓	✓
	Early-burst	∼	∼	✓	∼	✓
	Two-infall	✓	∼	✓	✓	✓

Notes. See discussion in Section 4.1 for the [Fe/H] DF, Section 4.2 for the [O/Fe] DF, Section 4.3 for the [O/Fe] bimodality, Section 4.4 for the [O/Fe]–[Fe/H] plane, and Section 4.5 for the age–[O/Fe] plane. Table 7 presents the quantitative scores used to make these comparisons.

SFH, while the exponential DTD with $\tau = 3$ Gyr produces better agreement with the data for that SFH. Finally, the inside-out SFH generally does not reproduce the APOGEE age–[O/Fe] relation well, but it scores better than average when combined with the triple-system DTD.

Our model scores are highly sensitive to small changes in the nucleosynthetic yields. A decrease in the SN Ia yield of Fe to $y_{\rm Fe}^{\rm Ia} = 0.0017$, which shifts the end point of the gas abundance tracks up by $\sim +0.05$ dex in [O/Fe], produces dramatically different scores for many of the models. This is because the KL divergence tests penalize distributions that are not well aligned with the data, even if the general trends and shape of the distribution are reproduced. For example, if the two-infall models are run with $y_{\rm Fe}^{\rm Ia} = 0.0017$, the abundance tracks do not dip below solar [O/Fe] (see the bottom-right panel of Figure 9), and consequently they outscore every other SFH. Small adjustments in the yields can affect the quality of the fit between our models and the data, so we caution against overinterpreting the qualitative comparisons in Table 6.

We also run a two-sample Kolmogorov–Smirnov (KS) test on the model [Fe/H] and [O/Fe] DFs to estimate the significance of the agreement with the observed distributions. However, we consistently reject the null hypothesis that the model and observed abundances are drawn from the same distribution at very high significance ($p << 0.05$) in each Galactic region. The large sample size means that even small deviations from the APOGEE distribution result in a very small p-value, making the KS test a poor diagnostic for model comparison. Overall, this reinforces the conclusion that even our best-performing models cannot reproduce all observations.

5.2. The Two-infall SFH

There have been many comparative GCE studies of the DTD with the two-infall model, providing an important point of comparison with our models. For example, Matteucci et al. (2006) explored the consequences of the two-population DTD (Mannucci et al. 2006), finding that its very high prompt SN Ia rate began to pollute the ISM during the halo phase and led to a faster decline in [O/Fe] with [Fe/H]. Matteucci et al. (2009) compared several DTDs, including the analytic forms of Greggio (2005) and the two-population DTD, in a multizone GCE model of the disk. Their comparisons to data were limited

to the solar neighborhood, and unlike our models, they did not factor in radial migration or gas outflows. Nevertheless, their conclusions align fairly well with ours: a relatively low fraction of prompt SNe Ia is needed to produce good agreement with observations.

More recently, Palicio et al. (2023) compared a similar suite of DTDs in one-zone models with a two-infall SFH. In contrast to previous studies of the two-infall model (e.g., Chiappini et al. 1997; Matteucci et al. 2009; Spitoni et al. 2021), they did incorporate gas outflows, making their models especially well suited to compare to ours. By modifying their yields, outflow mass-loading factor, and some of the parameters of their SFH, Palicio et al. (2023) were able to achieve a good fit to solar neighborhood abundance data for both the SD and DD analytic DTDs, which are approximated by our exponential ($\tau = 1.5$ Gyr) and plateau ($W = 1$ Gyr) models, respectively. Our results and theirs highlight the need for independent constraints on the SFH to resolve degeneracies with the DTD.

To our knowledge, this paper is the first exploration of the two-infall SFH in a multizone GCE model that incorporates both mass-loaded outflows and radial migration. A detailed examination of the parameters of the two-infall model is beyond the scope of this paper but will be the subject of future work.

5.3. Extragalactic Constraints

The power-law ($\alpha = -1.1$) DTD has the strongest observational motivation but poorly reproduces the disk abundance distributions. This can be mitigated somewhat with a longer minimum delay time, which has an effect on chemical evolution tracks similar to that of the addition of an initial plateau in the DTD (see discussion in Section 3). Even so, it is clear that the high fraction of prompt SNe Ia in extragalactic constraints on the DTD by, e.g., Maoz & Graur (2017) is at odds with Galactic chemical abundance measurements.

This tension could suggest that the Milky Way obeys a different DTD than other galaxies. This would not be too far beyond the finding of Maoz & Graur (2017) that field galaxies and galaxy clusters have a different DTD slope. However, Walcher et al. (2016) argued that the similarity of the age–[α/Fe] relation between solar neighborhood stars and nearby elliptical galaxies is evidence for a universal DTD. A physical mechanism would be needed to produce a different slope or form for the DTD in different environments, such as a metallicity dependence in the fraction of close binaries (e.g., Moe et al. 2019).

On the other hand, the difference between constraints from GCE models and extragalactic surveys indicates that these types of studies are most sensitive to different regimes of the DTD. Our results demonstrate that the high-α sequence in GCE models is highly sensitive to the DTD at short delay times. Measurements of galactic or cosmic SFHs typically provide constraints for the DTD in coarse age bins, with especially large uncertainties in the youngest bins (e.g., Maoz & Mannucci 2012), and it is difficult to constrain the SFH of individual galaxies at long look-back times (Conroy 2013). Additionally, measurements of the cosmic SN Ia rate become considerably uncertain at $z \gtrsim 1$ (see, e.g., Palicio et al. 2024). As a result, constraints from external galaxies should be more sensitive to the DTD at long delay times.

Palicio et al. (2024) fit combinations of cosmic star formation rates (CSFRs) and DTDs, many of which are similar to the forms in this paper, to the observed cosmic SN Ia rate. Notably, the DTD that best fit the majority of their CSFRs was the single-degenerate DTD of Matteucci & Recchi (2001), which is similar to the exponential form with $\tau = 1.5$ Gyr (see Appendix B for more discussion). They were able to exclude DTDs with a very high or very low fraction of prompt SNe Ia, but a number of their DTDs could produce a convincing fit to the observed rates with the right CSFR. Despite a very different methodology, their results mirror ours: that many forms for the DTD can produce a reasonable fit to the data when combined with the right SFH.

6. Conclusions

We have explored the consequences of eight different forms for the SN Ia DTD in multizone GCE models with radial migration. For each DTD, we explored combinations with four different popular SFHs from the literature, which represent a broad range of behavior over the lifetime of the disk seen in many prior GCE models. We compared our model outputs to abundances from APOGEE and ages from L23 for stars across the Milky Way disk. For each model, we computed a numerical score that reflects the agreement between the predictions and data across the entire disk for five observables. Our main conclusions are as follows:

1. While some combinations of SFH and DTD perform better than others, none of our models are able to reproduce every observed feature of the Milky Way disk.
2. The plateau DTD with a width $W = 1$ Gyr is best able to reproduce the observed abundance patterns for three of the four SFHs. For the inside-out SFH, it is narrowly surpassed by the (similar) triple-system DTD.
3. In general, we favor a DTD with a small fraction of prompt SNe Ia. The models with exponential, plateau, and triple-system DTDs perform significantly better than the models with two-population and power-law DTDs across all four SFHs.
4. The observationally derived $t^{-1.1}$ power-law DTD produces too few high-α stars. This could be mitigated with a longer minimum delay time or the addition of an initial plateau in the DTD at short delay times.
5. The SFH is the critical factor for producing a bimodal [α/Fe] distribution at fixed [Fe/H]. On its own, the DTD cannot produce a bimodal [α/Fe] distribution that matches what is observed. However, it does affect the location and strength of the high-α sequence, potentially enhancing the [α/Fe] bimodality resulting from the choice of SFH.

The origin of the Milky Way's [α/Fe] bimodality remains disputed. Some authors have argued that the combination of inside-out growth and radial migration sufficiently explains the observed distribution (e.g., Kubryk et al. 2015; Sharma et al. 2021; Chen et al. 2023; Prantzos et al. 2023), while others have argued that multiple episodes of gas infall at early times are required (e.g., Chiappini et al. 1997; Mackereth et al. 2018; Spitoni et al. 2019, 2020, 2021). In agreement with J21, we find that a smooth SFH combined with radial migration does not suffice. We find that these parameter choices predict too many stars between the high- and low-α sequences, resulting in a broad but unimodal [α/Fe] distribution at fixed [Fe/H].

We found that the MDF is least able to provide constraints on the DTD. The MDF is more sensitive to the SFH, but

overall trends across the Galaxy are primarily driven by the assumed radial abundance gradient and stellar migration prescription. However, the MDF is more sensitive to the DTD in the inner Galaxy due to the more sharply declining SFH (see discussion in Section 4.1). Under the model of inside-out formation, the MDF in the inner Galaxy traces older populations that are more sensitive to the enrichment of prompt SNe Ia.

We implemented a stellar migration scheme that reproduces the abundance trends seen in the models of J21, but produces smoother abundance distributions. Our method is flexible and is not tied to the output of a single hydrodynamical simulation. In future work, we will explore the effect of the strength and speed of radial migration on GCE models.

Recent studies have shown that the high specific SN Ia rates observed in low-mass galaxies (e.g., Brown et al. 2019; Wiseman et al. 2021) can be explained by a metallicity-dependent rate of SNe Ia (Gandhi et al. 2022; Johnson et al. 2023b). A similar metallicity dependence has also been observed in the rate of CCSNe (Pessi et al. 2023). These previous investigations varied only the normalization in the DTD. Gandhi et al. (2022) take into account radial migration by construction through their use of the FIRE-2 simulations. An exploration in the context of multizone models would be an interesting direction for future work, as would variations in the DTD shape.

Our results indicate that the allowed range of parameter space in GCE models is still too broad to precisely constrain the DTD. Future constraints may come from the Legacy Survey of Space and Time (LSST) at the Vera Rubin Observatory (Ivezić et al. 2019), which is expected to observe several million SNe during its 10 yr run. On the other hand, the improved sample size of SDSS-V (Kollmeier et al. 2017) will enable future GCE studies to constrain both the Galactic SFH and the DTD at a higher confidence.

Acknowledgments

We thank Prof. David H. Weinberg and attendees of OSU's Galaxy Hour for many useful discussions over the course of this project. We also thank the anonymous reviewer for providing thoughtful and constructive comments on the manuscript.

L.O.D. and J.A.J. acknowledge support from National Science Foundation grant no. AST-2307621. J.A.J. and J.W. J. acknowledge support from National Science Foundation grant no. AST-1909841. L.O.D. acknowledges financial support from an Ohio State University Fellowship. J.W.J. acknowledges financial support from an Ohio State University Presidential Fellowship and a Carnegie Theoretical Astrophysics Center postdoctoral fellowship.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss4.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC; https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

We would like to acknowledge the land that The Ohio State University occupies is the ancestral and contemporary territory of the Shawnee, Potawatomi, Delaware, Miami, Peoria, Seneca, Wyandotte, Ojibwe, and many other Indigenous peoples. Specifically, the university resides on land ceded in the 1795 Treaty of Greeneville and the forced removal of tribes through the Indian Removal Act of 1830. As a land grant institution, we want to honor the resiliency of these tribal nations and recognize the historical contexts that has and continues to affect the Indigenous peoples of this land.

Software: `VICE` (Johnson & Weinberg 2020), Astropy (Astropy Collaboration et al. 2013, 2018, 2022), scikit-learn (Pedregosa et al. 2011), SciPy (Virtanen et al. 2020), Matplotlib (Hunter 2007).

Appendix A
Reproducibility

This study was carried out using the reproducibility software show your work! (Luger et al. 2021), which leverages continuous integration to programmatically download the data from https://zenodo.org, create the figures, and compile the manuscript. Each figure caption contains two links: one to the data set stored on Zenodo used in the corresponding figure, and the other to the script used to make the figure (at the commit corresponding to the current build of the manuscript). The Git repository associated with this study is publicly available at https://github.com/lodubay/galactic-dtd, and the release v1.1.1 allows anyone to rebuild the entire manuscript. The multizone model outputs and APOGEE sample data set are stored at doi:10.5281/zenodo.12207380, and the source code for this article is available from Dubay (2024).

Appendix B
Analytic DTDs

Greggio (2005) derived analytic DTDs for SD and DD progenitor systems from assumptions about binary stellar evolution and mass exchange. Significant parameters for the

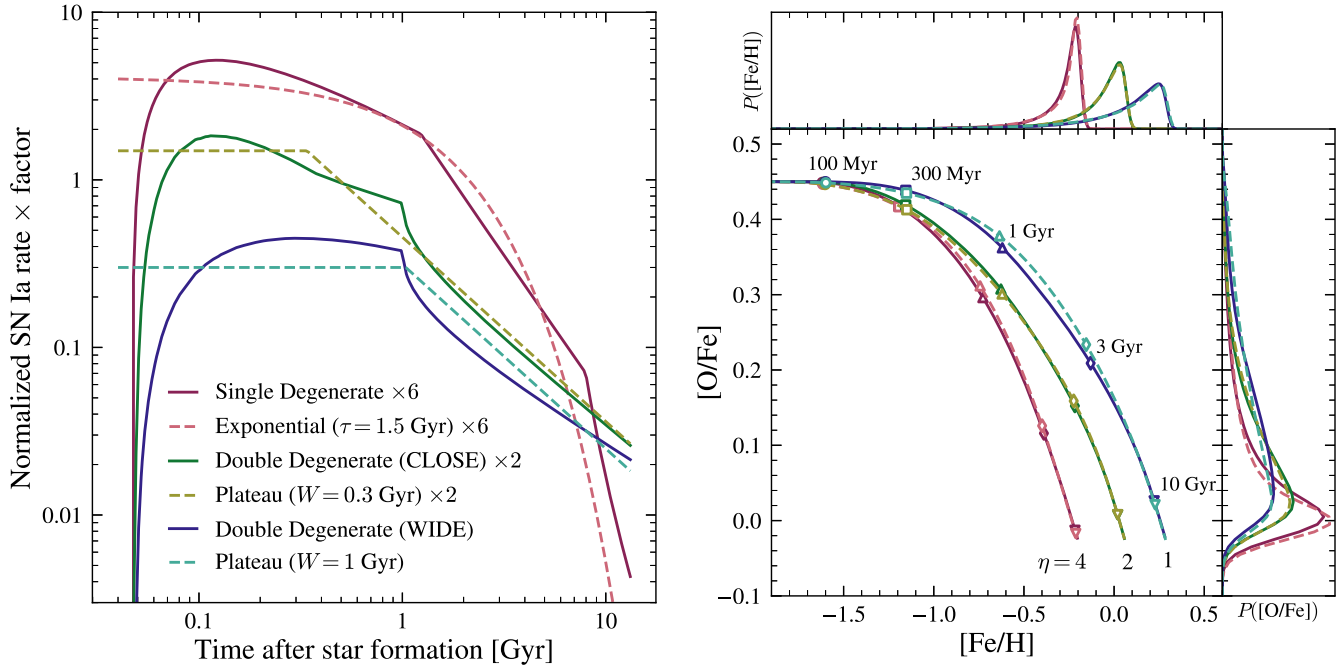

Figure 14. Left: Analytic DTDs from (Greggio 2005, solid curves) and simplified approximations thereof (dashed curves; see Section 2.2). Some functions are presented with a constant multiplicative factor for visual clarity. Right: Abundance tracks and distributions from one-zone models with the analytic and simple DTDs (same color scheme). For visual clarity, we vary the mass-loading factor to be $\eta = 4$, $\eta = 2$, and $\eta = 1$ for the red, green, and blue curves, respectively. All other model parameters are identical.

shape of the DTD are the distribution and range of stellar masses in progenitor systems, the efficiency of accretion in the SD scenario, and the distribution of separations at birth in the DD scenario. The left-hand panel of Figure 14 shows several of the Greggio (2005) analytic DTDs: one for SD progenitors, and two different prescriptions for DD progenitors ("WIDE" and "CLOSE"). The difference between the DD prescriptions relates to the ratio between the separation of the DD system and the initial separation of the binary, A/A_0. In the "WIDE" scheme, it is assumed that A/A_0 spans a wide distribution, and that the distributions of A and total mass of the system $m_{\rm DD}$ are independent, so one cannot necessarily predict the total merge time of a system based on its initial parameters. In the "CLOSE" scheme, there is assumed to be a narrow distribution of A/A_0 and a correlation between A and $m_{\rm DD}$, so the most massive binaries tend to merge quickly and the least massive merge last.

Here, we state our assumptions for the hyperparameters that can affect the shape of the Greggio (2005) DTDs. For the SD case, we assume a power-law slope of the mass ratio distribution with $\gamma = 1$, an efficiency of mass transfer $\epsilon = 1$, and maximum initial primary mass of $8\,M_\odot$. For the DD channel, we additionally assume a nuclear timescale for the least massive secondary $\tau_{\rm n,x} = 1$ Gyr, a minimum gravitational inspiral delay $\tau_{\rm gw,i} = 1$ Myr, an exponent of the power-law distribution of final separations $\beta_{\rm a} = 0$ (for the WIDE scheme), and an exponent of the power-law distribution of gravitational delays $\beta_{\rm g} = -0.75$ (for the CLOSE scheme).

In the left-hand panel of Figure 14, we also include simple functions that approximate the analytic DTDs of Greggio (2005). Chemical abundance distributions are sensitive to the broad shape of the DTD but are agnostic to the underlying physics of the progenitor systems. These simplified forms reduce the number of free parameters for the DTD and make the GCE model predictions easier to interpret.

The right-hand panel of Figure 14 shows the results of one-zone chemical evolution models with the Greggio (2005) DTDs and our simplified forms. We use the same model parameters as in Section 3 but with different values of η to spread the tracks out visually in [Fe/H]. The model with the SD DTD follows a nearly identical track to the exponential ($\tau = 1.5$ Gyr) DTD, and they produce very similar distributions of [O/Fe]. Likewise, the DD CLOSE DTD is well approximated by the plateau DTD with $W = 0.3$ Gyr and a power-law slope $\alpha = -1.1$. The WIDE prescription is also best approximated by a plateau DTD, but with a longer plateau width of $W = 1$ Gyr. In all cases, the effect of the difference between the analytic DTD and its simple approximation is too small to be observed. We also ran a multizone model with the inside-out SFH and the Greggio (2005) SD DTD and found it produced results that are nearly identical to those from the model with the exponential ($\tau = 1.5$ Gyr) DTD.

Appendix C
Stellar Migration

In their multizone models, J21 randomly assign an analog star particle from h277, adopting its radial migration distance ΔR and final midplane distance z, for each stellar population generated by VICE. The analogs are chosen such that the star particle was born at a similar radius and time as the stellar population in the GCE model. This prescription allows VICE to adopt a realistic pattern of radial migration without needing to implement its own hydrodynamical simulation. However, in regions where the number of h277 star particles is relatively low, such as at large $R_{\rm gal}$ and small t, a single h277 star particle can be assigned as an analog to multiple VICE stellar populations. These populations will have similar formation and migration histories and consequently similar abundances, which produces unphysical "clumps" of stars in the abundance distributions at large $|z|$ and $R_{\rm gal}$.

We adopt a prescription for radial migration that produces smoother abundance distributions while still following the behavior of h277. We fit a Gaussian to the distribution of $\Delta R = R_{\rm final} - R_{\rm form}$ from the h277 output, binned by both formation radius $R_{\rm form}$ and age. We are motivated by the findings of Okalidis et al. (2022) that the strength of stellar migration in the Auriga simulations (Grand et al. 2017) varies with both $R_{\rm form}$ and age. Each Gaussian is centered at 0, and we find that the scale $\sigma_{\rm RM}$ is best described by the function

$$\sigma_{\rm RM} = \sigma_{\rm RM8}\left(\frac{\tau}{8\ {\rm Gyr}}\right)^{0.33}\left(\frac{R_{\rm form}}{8\ {\rm kpc}}\right)^{0.61}, \tag{C1}$$

where τ is the stellar age and $\sigma_{\rm RM8} = 2.68$ kpc describes the migration strength for an 8 Gyr old population with $R_{\rm form} = 8$ kpc. For comparison, Frankel et al. (2018) found a steeper τ-dependence of $\sigma_{\rm RM} \propto \tau^{1/2}$ and a higher scaling of $\sigma_{\rm RM8} = 3.6$ kpc for a sample of APOGEE red clump stars. Our age scaling is in good agreement with Lu et al. (2024), who find that radial migration in galaxies from the NIHAO simulations (Wang et al. 2015) follow a relatively universal relation of $\sigma_{\rm RM} \propto \tau^{0.32}$, but with a slightly higher $\sigma_{\rm RM8} \approx 3$ kpc. We use the lower value here because it reproduces the trends in h277, and by extension J21, but the $\sim$25% difference may affect the predictions.

When VICE forms a stellar population at initial radius $R_{\rm form}$, we assign a value of ΔR by randomly sampling from a Gaussian with a width given by Equation (C1). The star particle migrates to its final radius $R_{\rm final}$ in a similar manner to the "diffusion" case from J21, but with a time dependence $\propto \Delta t^{1/3}$, motivated by the age scaling of $\sigma_{\rm RM}$.

We note that the h277 galaxy has a weak and transient bar, in contrast to the Milky Way. The presence of a strong bar has been found to affect the strength of radial migration throughout the disk (e.g., Brunetti et al. 2011) and lead to a flattening of the metallicity gradient for old populations (Okalidis et al. 2022).

Figure 15 compares the distributions of $R_{\rm final}$ in bins of $R_{\rm form}$ and stellar age between the h277 analog method and our new prescription. There is good agreement across the Galaxy in the youngest age bins, but the "clumpiness" of the h277 analog populations, a consequence of sampling noise, becomes evident for old stars formed in the outer Galaxy. The distribution of h277 star particles in the $10 \leqslant \tau < 12$ Gyr and $11 \leqslant R_{\rm form} < 13$ kpc bin indicates significant inward migration due to a merging satellite. Our Gaussian sampling scheme eliminates both the clumpiness and the impact of mergers and other external events on radial migration.

Like J21, we assume all stellar populations form in the midplane ($z = 0$). J21 take the present-day midplane distance $z_{\rm final}$ directly from the h277 analog particle. To produce smoother abundance distributions, we fit a sech^2 function (Spitzer 1942) to the distribution of z in h277. The PDF of

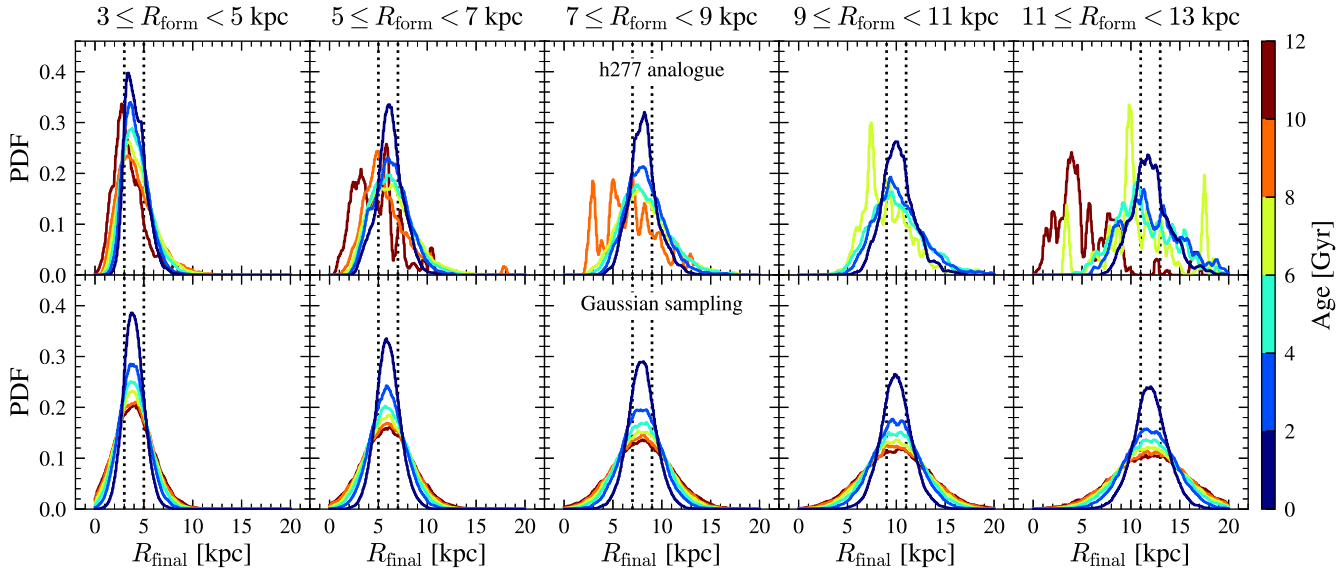

Figure 15. The distribution of final radius $R_{\rm final}$ as a function of formation radius $R_{\rm form}$ and age for the h277 analog (top row) and Gaussian sampling scheme (bottom row; see discussion in Appendix C). From left to right, star particles are binned by formation annulus, as noted at the top of each column of panels. Within each panel, colored curves represent the different age bins, ranging from the youngest stars (dark blue) to the oldest (dark red). In the top row, we exclude age bins with fewer than 100 unique analog IDs for visual clarity. All distributions are normalized so that the area under the curve is 1, and they have been boxcar-smoothed with a window width of 0.5 kpc. The vertical dotted black lines indicate the bounds of each bin in $R_{\rm form}$; stars within that region of the distribution have not migrated significantly far from their birth radius.

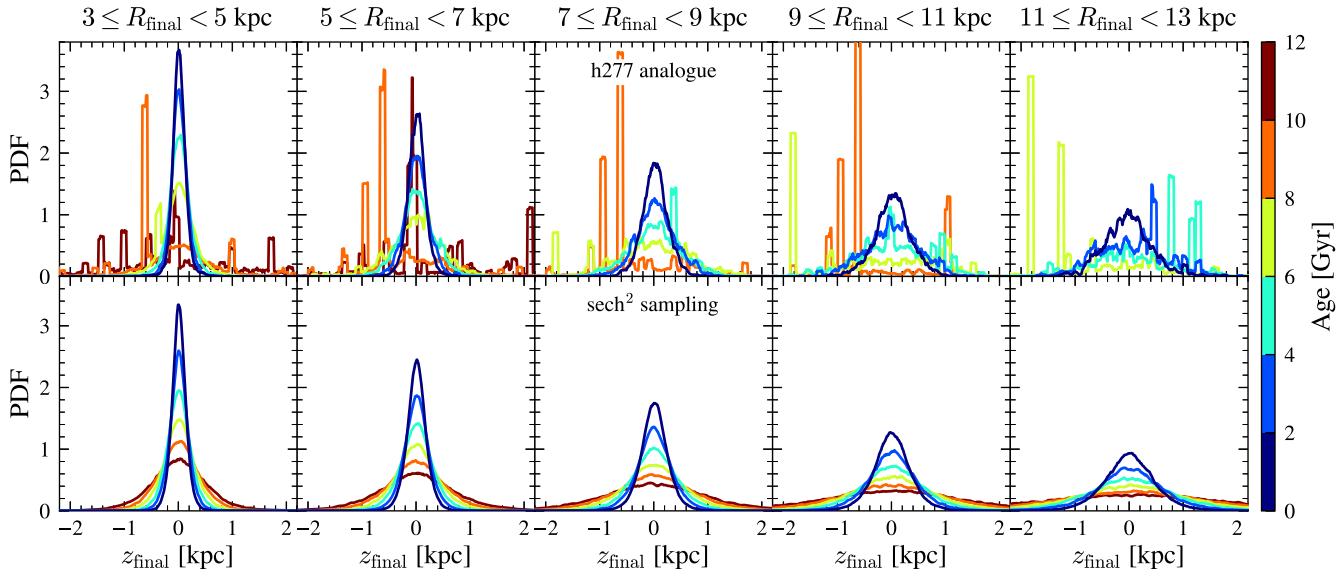

Figure 16. Similar to Figure 15, but for the distribution of present-day midplane distance $z_{\rm final}$ as a function of radius and age. From left to right, star particles are binned by *final* annulus. In the top row, we exclude age bins with fewer than 500 unique analog IDs for clarity. All distributions have been boxcar-smoothed with a window width of 0.1 kpc.

$z_{\rm final}$ given some scale height h_z is

$$\mathrm{PDF}(z_{\rm final}) = \frac{1}{4h_z}\,\mathrm{sech}^2\left(\frac{z_{\rm final}}{2h_z}\right). \tag{C2}$$

We fit Equation (C2) to the distributions of z in `h277` in varying bins of T and $R_{\rm final}$. We find that h_z is best described by the function

$$h_z = (h_{z,s}/e^2)\exp(\tau/\tau_s + R_{\rm final}/R_s), \tag{C3}$$

where $h_{z,s} = 0.24$ kpc is the scale height at $\tau_s = 7$ Gyr and $R_s = 6$ kpc. For each star particle in `VICE`, we sample $z_{\rm final}$ from the distribution described by Equation (C2) with a width given by Equation (C3). Figure 16 shows the resulting distributions of $z_{\rm final}$ are similar to the `h277` analog scheme for all but the oldest stellar populations.

Appendix D
Quantitative Comparison Scores

Table 7 presents the quantitative scores that measure the difference between the multizone outputs and APOGEE data for four observables. Details of the calculations for each observable are presented in the corresponding subsections of Section 4, but in summary, divergence statistics between the multizone output and APOGEE data (Equation (18) for the [Fe/H] and [O/Fe] DFs, Equation (20) for the [O/Fe]–[Fe/H] plane, and Equation (21) for the age–[O/Fe] plane) are computed within each Galactic region. The overall score for the model is the average of the divergence statistics across all regions, weighted by the number of APOGEE targets in each region. The scores can be strongly affected by zero-point offsets between the model and observed abundance distributions, so we caution against using these scores as an absolute metric of model performance.

Table 7
Quantitative Scores Comparing the Model Output, APOGEE DR17 Abundances, and L23 Ages for Each Multizone Model

DTD	SFH	MDF	[O/Fe] DF	[Fe/H]–[O/Fe]	Age–[O/Fe]
Two-population ($t_p = 0.05$ Gyr)	Inside-out	0.292	0.947	1.992	3.41
	Late-burst	0.291	0.99	1.996	3.069
	Early-burst	0.211	1.591	2.6	2.439
	Two-infall	0.219	0.638	1.839	2.873
Power-law ($\alpha = -1.4$)	Inside-out	0.343	1.318	2.695	4.211
	Late-burst	0.327	1.35	2.73	3.608
	Early-burst	0.233	1.73	3.424	3.084
	Two-infall	0.262	0.898	2.46	3.498
Power-law ($\alpha = -1.1$)	Inside-out	0.288	0.688	1.513	3.255
	Late-burst	0.292	0.731	1.552	2.956
	Early-burst	0.215	1.197	2.471	2.257
	Two-infall	0.202	0.437	1.327	2.56
Exponential ($\tau = 1.5$ Gyr)	Inside-out	0.24	0.588	1.119	3.158
	Late-burst	0.222	0.388	0.921	2.54
	Early-burst	0.247	0.954	2.027	2.224
	Two-infall	0.18	0.482	1.156	2.283
Exponential ($\tau = 3.0$ Gyr)	Inside-out	0.183	0.328	0.773	1.976
	Late-burst	0.214	0.253	0.682	1.788
	Early-burst	0.217	0.732	1.391	1.52
	Two-infall	0.134	0.484	0.897	1.518
Plateau ($W = 0.3$ Gyr)	Inside-out	0.235	0.353	0.837	2.509
	Late-burst	0.256	0.372	0.854	2.241
	Early-burst	0.215	0.663	1.656	1.731
	Two-infall	0.154	0.349	0.863	1.845
Plateau ($W = 1.0$ Gyr)	Inside-out	0.199	0.243	0.679	1.844
	Late-burst	0.243	0.314	0.73	1.798
	Early-burst	0.211	0.567	1.207	1.399
	Two-infall	0.126	0.422	0.799	1.37
Triple-system ($t_{\rm rise} = 0.5$ Gyr)	Inside-out	0.179	0.231	0.678	1.568
	Late-burst	0.248	0.327	0.758	1.569
	Early-burst	0.223	0.603	1.239	1.339
	Two-infall	0.119	0.481	0.812	1.281

Notes. See discussion in Section 4.1 for the [Fe/H] DF, Section 4.2 for the [O/Fe] DF, Section 4.4 for the [O/Fe]–[Fe/H] plane, and Section 4.5 for the age–[O/Fe] plane. Table 6 summarizes the relative performance of each model based on the scores presented here.

ORCID iDs

Liam O. Dubay https://orcid.org/0000-0003-3781-0747
Jennifer A. Johnson https://orcid.org/0000-0001-7258-1834
James W. Johnson https://orcid.org/0000-0002-6534-8783

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Andrews, B. H., Weinberg, D. H., Schönrich, R., & Johnson, J. A. 2017, ApJ, 835, 224
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L, et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Bennett, M., & Bovy, J. 2019, MNRAS, 482, 1417
Bensby, T., Feltzing, S., & Oey, M. S. 2014, A&A, 562, A71
Benz, W., Hills, J. G., & Thielemann, F. K. 1989, ApJ, 342, 986
Bilitewski, T., & Schönrich, R. 2012, MNRAS, 426, 2266
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Bowen, I. S., & Vaughan, A. H. J. 1973, ApOpt, 12, 1430
Brown, J. S., Stanek, K. Z., Holoien, T. W. S., et al. 2019, MNRAS, 484, 3785
Brunetti, M., Chiappini, C., & Pfenniger, D. 2011, A&A, 534, A75
Castrillo, A., Ascasibar, Y., Galbany, L., et al. 2021, MNRAS, 501, 3122
Chen, B., Hayden, M. R., Sharma, S., et al. 2023, MNRAS, 523, 3791
Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765
Chieffi, A., & Limongi, M. 2004, ApJ, 608, 405
Chomiuk, L., Soderberg, A. M., Chevalier, R. A., et al. 2016, ApJ, 821, 119
Christensen, C., Quinn, T., Governato, F., et al. 2012, MNRAS, 425, 3058
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., Bonaca, A., Cargile, P., et al. 2019, ApJ, 883, 107
Conroy, C., Weinberg, D. H., Naidu, R. P., et al. 2022, arXiv:2204.02989
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
Do, A., Shappee, B. J., De Cuyper, J.-P., et al. 2021, MNRAS, 508, 3649
Dubay, L. O., Tucker, M. A., Do, A., Shappee, B. J., & Anand, G. S. 2022, ApJ, 926, 98
Dubay, L. 2024, lodubay/galactic-dtd: Source Code for Dubay et al. (2024), Zenodo, doi:10.5281/zenodo.12521399
Edvardsson, B., Andersen, J., Gustafsson, B., et al. 1993, A&A, 275, 101
Fang, X., Thompson, T. A., & Hirata, C. M. 2018, MNRAS, 476, 4234
Fausnaugh, M. M., Vallely, P. J., Kochanek, C. S., et al. 2021, ApJ, 908, 51
Feuillet, D. K., Frankel, N., Lind, K., et al. 2019, MNRAS, 489, 1742
Förster, F., Wolf, C., Podsiadlowski, P., & Han, Z. 2006, MNRAS, 368, 1893
Frankel, N., Rix, H.-W., Ting, Y.-S., Ness, M., & Hogg, D. W. 2018, ApJ, 865, 96
Freundlich, J., & Maoz, D. 2021, MNRAS, 502, 5882
Friedmann, M., & Maoz, D. 2018, MNRAS, 479, 3563
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gandhi, P. J., Wetzel, A., Hopkins, P. F., et al. 2022, MNRAS, 516, 1941
García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144
Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179
Graur, O., Bianco, F. B., & Modjaz, M. 2015, MNRAS, 450, 905
Graur, O., & Maoz, D. 2013, MNRAS, 430, 1746
Graur, O., Rodney, S. A., Maoz, D., et al. 2014, ApJ, 783, 28
GRAVITY Collaboration, Abuter, R., Amorim, A., et al. 2018, A&A, 615, L15
Greggio, L. 2005, A&A, 441, 1055
Greggio, L., Renzini, A., & Daddi, E. 2008, MNRAS, 388, 829
Grisoni, V., Chiappini, C., Miglio, A., et al. 2024, A&A, 683, A111
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, ApJ, 808, 132
Hekker, S., & Johnson, J. A. 2019, MNRAS, 487, 4343
Heringer, E., Pritchet, C., & van Kerkwijk, M. H. 2019, ApJ, 882, 52
Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, AJ, 150, 148
Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800
Hunter, J. D. 2007, CSE, 9, 90
Iben, I. J., & Tutukov, A. V. 1984, ApJS, 54, 335
Isern, J. 2019, ApJL, 878, L11
Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, ApJS, 125, 439
Johnson, J. W., Conroy, C., Johnson, B. D., et al. 2023a, MNRAS, 526, 5084
Johnson, J. W., Kochanek, C. S., & Stanek, K. Z. 2023b, MNRAS, 526, 5911
Johnson, J. W., & Weinberg, D. H. 2020, MNRAS, 498, 1364
Johnson, J. W., Weinberg, D. H., Vincenzo, F., et al. 2021, MNRAS, 508, 4484
Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, AJ, 160, 120
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv:1711.03234
Kubryk, M., Prantzos, N., & Athanassoula, E. 2015, A&A, 580, A126
Kullback, S., & Leibler, R. A. 1951, Ann. Math. Stat., 22, 79
Lacey, C. G., & Fall, S. M. 1985, ApJ, 290, 154
Leung, H. W., Bovy, J., Mackereth, J. T., & Miglio, A. 2023, MNRAS, 522, 4577
Licquia, T. C., & Newman, J. A. 2015, ApJ, 806, 96
Limongi, M., & Chieffi, A. 2006, ApJ, 647, 483
Liu, Z.-W., Röpke, F. K., & Han, Z. 2023, RAA, 23, 082001
Livio, M., & Mazzali, P. 2018, PhR, 736, 1
Lu, Y., Buck, T., Nidever, D., et al. 2024, MNRAS, 532, 411
Luger, R., Bedell, M., Foreman-Mackey, D., et al. 2021, arXiv:2110.06271
Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, MNRAS, 489, 176
Mackereth, J. T., Crain, R. A., Schiavon, R. P., et al. 2018, MNRAS, 477, 5072

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Mannucci, F., Della Valle, M., Panagia, N., et al. 2005, A&A, 433, 807
Mannucci, F., Della Valle, M., & Panagia, N. 2006, MNRAS, 370, 773
Maoz, D., & Graur, O. 2017, ApJ, 848, 25
Maoz, D., & Mannucci, F. 2012, PASA, 29, 447
Maoz, D., Mannucci, F., & Brandt, T. D. 2012, MNRAS, 426, 3282
Maoz, D., Mannucci, F., & Nelemans, G. 2014, ARA&A, 52, 107
Maoz, D., Sharon, K., & Gal-Yam, A. 2010, ApJ, 722, 1879
Martig, M., Fouesneau, M., Rix, H.-W., et al. 2016, MNRAS, 456, 3655
Matteucci, F., & Greggio, L. 1986, A&A, 154, 279
Matteucci, F., Panagia, N., Pipino, A., et al. 2006, MNRAS, 372, 265
Matteucci, F., & Recchi, S. 2001, ApJ, 558, 351
Matteucci, F., Spitoni, E., Recchi, S., & Valiante, R. 2009, A&A, 501, 531
Maza, J., & van den Bergh, S. 1976, ApJ, 204, 519
Moe, M., Kratter, K. M., & Badenes, C. 2019, ApJ, 875, 61
Mor, R., Robin, A. C., Figueras, F., Roca-Fàbrega, S., & Luri, X. 2019, A&A, 624, L1
Nidever, D. L., Holtzman, J. A., Allende Prieto, C., et al. 2015, AJ, 150, 173
Nomoto, K. 1982, ApJ, 253, 798
Nomoto, K., & Iben, I. J. 1985, ApJ, 297, 531
Okalidis, P., Grand, R. J. J., Yates, R. M., & Springel, V. 2022, MNRAS, 514, 5085
Pakmor, R., Kromer, M., Taubenberger, S., et al. 2012, ApJL, 747, L10
Palicio, P. A., Matteucci, F., Della Valle, M., & Spitoni, E. 2024, arXiv:2402.16635
Palicio, P. A., Spitoni, E., Recio-Blanco, A., et al. 2023, A&A, 678, A61
Palla, M. 2021, MNRAS, 503, 3216
Panagia, N., Van Dyk, S. D., Weiler, K. W., et al. 2006, ApJ, 646, 369
Pantoni, L., Lapi, A., Massardi, M., Goswami, S., & Danese, L. 2019, ApJ, 880, 129
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
Perez-Cruz, F. 2008, in 2008 IEEE Int. Symp. on Information Theory (Piscataway, NJ: IEEE), 1666
Pessi, T., Anderson, J. P., Lyman, J. D., et al. 2023, ApJL, 955, L29
Poulhazan, P.-A., Scannapieco, C., & Creasey, P. 2018, MNRAS, 480, 4817
Prantzos, N., Abia, C., Chen, T., et al. 2023, MNRAS, 523, 2126
Rajamuthukumar, A. S., Hamers, A. S., Neunteufel, P., Pakmor, R., & de Mink, S. E. 2023, ApJ, 950, 9
Rebassa-Mansergas, A., Toonen, S., Korol, V., & Torres, S. 2019, MNRAS, 482, 3656
Rodney, S. A., Riess, A. G., Strolger, L.-G., et al. 2014, AJ, 148, 13
Ruiter, A. J. 2020, IAUS, 357, 1
Rybizki, J., Just, A., & Rix, H.-W. 2017, A&A, 605, A59
Sahlholdt, C. L., Feltzing, S., & Feuillet, D. K. 2022, MNRAS, 510, 4669
Saio, H., & Nomoto, K. 1998, ApJ, 500, 388
Sánchez, S. F. 2020, ARA&A, 58, 99
Sandford, N. R., Weinberg, D. H., Weisz, D. R., & Fu, S. W. 2024, MNRAS, 530, 2315
Scannapieco, E., & Bildsten, L. 2005, ApJL, 629, L85
Schaefer, B. E., & Pagnotta, A. 2012, Natur, 481, 164
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Schönrich, R., & Binney, J. 2009, MNRAS, 396, 203
Sharma, S., Hayden, M. R., & Bland-Hawthorn, J. 2021, MNRAS, 507, 5882
Shen, K. J., Bildsten, L., Kasen, D., & Quataert, E. 2012, ApJ, 748, 35
Silva Aguirre, V., Bojsen-Hansen, M., Slumstrup, D., et al. 2018, MNRAS, 475, 5487
Simonetti, P., Matteucci, F., Greggio, L., & Cescutti, G. 2019, MNRAS, 486, 2896
Spitoni, E., Silva Aguirre, V., Matteucci, F., Calura, F., & Grisoni, V. 2019, A&A, 623, A60
Spitoni, E., Verma, K., Silva Aguirre, V., & Calura, F. 2020, A&A, 635, A58
Spitoni, E., Verma, K., Silva Aguirre, V., et al. 2021, A&A, 647, A73
Spitzer, L. J. 1942, ApJ, 95, 329
Strolger, L.-G., Riess, A. G., Dahlen, T., et al. 2004, ApJ, 613, 200
Strolger, L.-G., Rodney, S. A., Pacifici, C., Narayan, G., & Graur, O. 2020, ApJ, 890, 140
Thompson, T. A. 2011, ApJ, 741, 82
Tinsley, B. M. 1979, ApJ, 229, 1046
Totani, T., Morokuma, T., Oda, T., Doi, M., & Yasuda, N. 2008, PASJ, 60, 1327
Tucker, M. A., & Shappee, B. J. 2024, ApJ, 962, 74
Tucker, M. A., Shappee, B. J., Vallely, P. J., et al. 2020, MNRAS, 493, 1044
Vincenzo, F., Matteucci, F., & Spitoni, E. 2017, MNRAS, 466, 2939
Vincenzo, F., Weinberg, D. H., Miglio, A., Lane, R. R., & Roman-Lopes, A. 2021, MNRAS, 508, 5903
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Walcher, C. J., Yates, R. M., Minchev, I., et al. 2016, A&A, 594, A61
Wallerstein, G. 1962, ApJS, 6, 407
Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, MNRAS, 454, 83
Webbink, R. F. 1984, ApJ, 277, 355
Weinberg, D. H., Andrews, B. H., & Freudenburg, J. 2017, ApJ, 837, 183
Weinberg, D. H., Griffith, E. J., Johnson, J. W., & Thompson, T. A. 2023, arXiv:2309.05719
Whelan, J., & Iben, I. J. 1973, ApJ, 186, 1007
Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001
Wiseman, P., Sullivan, M., Smith, M., et al. 2021, MNRAS, 506, 3330
Yong, D., Casagrande, L., Venn, K. A., et al. 2016, MNRAS, 459, 487
Yoon, S. C., & Langer, N. 2003, A&A, 412, L53