Title

Chemical Evolution Across Cosmic Time: Stellar Elemental Abundance Patterns and Radial Redistribution in Cosmological Simulations

Author

Bellardini, Matthew

Publication Date

2023

Peer reviewed|Thesis/dissertation

Chemical Evolution Across Cosmic Time: Stellar Elemental Abundance Patterns and Radial Redistribution in Cosmological Simulations

By

Matthew Alfred Bellardini

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Professor Andrew Wetzel

Professor Tucker Jones

Professor Stefano Valenti

Committee in Charge

2023

Contents

Abstract v

Acknowledgments vii

Chapter 1. Introduction 1

1.1. A Brief Overview of the Universe 1

1.2. Star Formation and Galactic Chemical Enrichment 2

1.3. Galactic Archaeology 3

1.4. Chemical Tagging 4

1.5. Understanding Galaxy Evolution with Cosmological Simulations 6

1.6. Dissertation Overview 7

Chapter 2. 3-D gas-phase elemental abundances across the formation histories of Milky Way-mass galaxies in the FIRE simulations: initial conditions for chemical tagging 8

2.1. Abstract 8

2.2. Introduction 9

2.3. Methods 15

2.4. Results 19

2.5. Summary and Discussion 49

Acknowledgements 58

Data Availability 59

Chapter 3. 3D elemental abundances of stars at formation across the histories of Milky Way-mass galaxies in the FIRE simulations 60
3.1. Abstract 60
3.2. Introduction 61
3.3. Methods 65
3.4. Results 71
3.5. Summary and Discussion 100
Acknowledgements 108
Data Availability 108

Chapter 4. Measuring stellar motion: Radial and angular momentum redistribution of stellar orbits in FIRE simulations of Milky Way-mass galaxies with varying definitions of orbital radius 109
4.1. Abstract 109
4.2. Introduction 110
4.3. Methods 112
4.4. Results 119
4.5. Summary and Discussion 140
Acknowledgements 145
Data Availability 145

Chapter 5. Conclusion 146
5.1. Summary 146
5.2. Future Work 148

Appendix A. Chapter 2 Appendices 152
A.1. Scaled radial profiles 152
A.2. All gas versus star-forming gas 153

A.3. Impact of diffusion coefficient 156

Appendix B. Chapter 3 Appendices 160

B.1. Shapes of abundance profiles 160

B.2. Stars at formation versus gas 162

B.3. Measurable homogeneity 163

Appendix C. Chapter 4 Appendices 167

C.1. Effect of Radial Selection 167

C.2. Binning stars by physical radius 168

C.3. Selecting by Formation Radius 170

Bibliography 175

Abstract

Understanding galaxy formation and evolution requires characterizing elemental abundance distributions in galaxies. Chemical tagging is a useful tool to understand the evolutionary history of the Milky Way (MW) because it takes advantage of the fact that stellar abundances at present-day are identical to the abundances with which stars formed. Thus, stellar elemental abundances provide an observable that can, in principle, be used to determine the birth location of a star.

Elemental abundance observations of gas and stars in the MW and nearby galaxies show that the median elemental abundance typically decreases with increasing radius with little scatter about the mean at each radius. These observations are robust for nearby galaxies, however they are less certain at high redshift because of the observational difficulties associated with obtaining spatially resolved spectra of high redshift galaxies. As a result, many galactic elemental evolution models rely on assumptions that observed properties of the MW and nearby external galaxies, crucially the lack of azimuthal scatter in abundances, are a time-independent property. Alternatively, some researchers use physical models of galaxy evolution to which they fit a multitude of free parameters such that their model recreates the present-day observed properties of the MW and external galaxies. However, these models typically rely on overly-simplified assumptions of physical processes and include multiple free functions unconstrained by physical models.

Cosmological zoom-in hydrodynamic simulations can help to precisely characterize the spatial distribution of stellar elemental abundances in MW-mass galaxies across cosmic time. Our results challenge the status quo of galactic elemental evolution models. We find that the minimal azimuthal abundance variations in MW-mass galaxies are not an intrinsic property, rather galaxies evolve from an epoch of extreme azimuthal abundance variation to their present-day state. Additionally, radial abundance gradients in galaxies were nearly non-existent at sufficiently high redshifts, despite being relatively strong at present-day.

These results suggest a higher degree of difficulty in chemically tagging stars than previously assumed. Thus, chemical tagging techniques may only be able to loosely constrain birth locations of older stellar populations. To help address this, we characterize the scale of radial redistribution of stars in simulated MW-mass galaxies as a function of both stellar age and stellar location. Accounting for stellar radial redistribution as a function of present-day radial location can help break degeneracies of birth location for older stellar populations. Our results on stellar radial redistribution suggest that inferring a time-dependent radial abundance gradient of the MW from present-day observations is non-trivial; at present-day, old stellar populations are often at very different radii than the radii at which they formed. Therefore, more emphasis must be placed on models derived from cosmological simulations.

Acknowledgments

If I were to individually thank everyone who has helped me along the way to accomplishing this dissertation, the acknowledgements would be longer than the dissertation. I can, unfortunately, only thank a subset of these people, but know that there are many more who made this possible.

To my advisor, Andrew Wetzel, you have been the epitome of an academic advisor. Throughout my graduate career you encouraged my growth not just as a scientist, but as a person. I can not thank you enough for your understanding when things have been difficult, your constant availability to help me with my every problem, and your unwavering encouragement. I could not have asked for a better advisor.

To Sarah Loebman, you are one of the most helpful people I have ever met. You have the unique ability to provide an easily understandable answer to any question while never being condescending. While you were only at Davis for a short time, I learned more from you than I thought possible. Also, I want to thank you tremendously for your support in networking with other scientists, especially as a young graduate student. I would not have the confidence I have now without your support.

To my research group, I am lucky to have been part of such a close knit collaboration. Thank you for your support throughout my graduate career. You have always been the people I can turn to when I have questions I feel embarrassed to ask elsewhere. I am grateful for all of the learning we did together.

To my office mates, Isaiah Santistevan, Pratik Gandhi, and Keerthi Vasan. Thank you for all of the discussion, both scientific and causal, and for constantly helping me ponder through research questions. To my friends in the department, thank you for the regular lunches and coffee breaks. They served as much needed deviations from the grind of research and were crucial in getting me out of the department.

To my undergraduate advisor, Luke Keller, you helped develop me as a scientist by introducing me to research and supporting me throughout my undergraduate years. To all of my undergraduate physics faculty, thank you for setting me on the path to succeed. The knowledge you taught me helped me deal with the coursework that inundated me in my first year of graduate school.

To my high school physics teacher, Jim Overhiser, you started my interest in physics. I am so grateful to have had you as a teacher as your passion and teaching style was the only reason I knew what I wanted to study going forward.

To my steadfast friends Kevin Flanagan, Devon Baris, and Elijah Cronk, you provided me with a welcome respite from graduate school. Our weekly meetups on zoom kept me sane during the most difficult times.

To Tami, Steve, Joleen, and Jordan your support and encouragement has been unwavering. Thank you for everything from providing me a place to stay to providing me with a family Christmas during the first year of the pandemic. You made California a home for me.

To my parents, I could not have done this without you. You have always encouraged me to follow my passions, you pushed me to always do my best, and you made me into who I am today. Words can not adequately express the thanks I have for you. Your unceasing support built up my confidence and the values you instilled in me gave me the resolve and determination to succeed.

To my fiancée, Raeanne, you deserve an award for what you have dealt with. Through all of the ups and downs of graduate school you have been by my side. You comforted me as I anxiously prepped for my qualifying exam. You provided me with the motivation I needed whenever I was struggling with a paper or a proposal. You celebrated my accomplishments and you helped me overcome my challenges. From the bottom of heart, thank you. Of course, I could not thank you without also thanking our cats Kelvin and Joule and our dog

Plum. While they often seemed to make things harder rather than easier, they provided much needed distractions.

CHAPTER 1

Introduction

1.1. A Brief Overview of the Universe

The evolution of the universe and its components is one of the most fundamental questions humanity can tackle. The current theoretical model of the universe begins with the big bang. The first few minutes of the universe involved an extremely energetic expansion of matter and radiation eventually leading to a state of coupled photons and charged particles. For the first ≈50,000 years of the universe, radiation was the dominant source of energy density (Ryden, 2017). After this, the universe had sufficiently expanded and cooled such that matter was the leading source of energy density. During this era of the universe, ≈370,000 years after the big bang, the universe cooled to the point that photons decoupled from charged particles so electrons and protons could bond forming neutral hydrogen and helium (e.g. Tanabashi et al., 2018).

After the decoupling of photons and charged particles matter overdensities in the universe seeded during inflation began to collapse (Press & Schechter, 1974). Hierarchical structure formation occurred through mergers of initially small dark matter halos (e.g. Ostriker & Tremaine, 1975) into larger dark matter halos within which galaxies formed (Searle & Zinn, 1978). This hierarchy of galaxy formation left observable imprints on the universe today, such as the prevalence of low-mass galaxies relative to higher mass galaxies (e.g. White & Rees, 1978; Kauffmann et al., 1993; Cole et al., 1994). Once these over densities are sufficiently massive to accrete gas, star formation can occur leading to stellar feedback processes which are crucial role to galaxy evolution (e.g. Vogelsberger et al., 2014; Wang et al., 2015; Hopkins et al., 2018).

1.2. Star Formation and Galactic Chemical Enrichment

As giant molecular clouds collapse under self-gravity, they fragment, and eventually these collapsing regions surpass a critical density and temperature threshold which triggers the fusion of hydrogen into helium. The energy released via this fusion counteracts the gravitational collapse of the cloud and thus a main sequence star is born. As nuclear fusion exhausts the hydrogen in the core of a star, radiation pressure no longer counteracts gravitational collapse leading to further contraction of the star. This collapses triggers hydrogen to helium fusion in a layer around a helium core. The helium core collapse continues until helium fusion is triggered. Through this constant fight between gravitational collapse and stellar nucleosynthesis-induced radiation pressure heavy elements are produced.

When stars eventually die in energetic supernovae events, they inject massive amounts of energy and momentum into the surrounding gas, as well as the metals that they formed throughout their lives (and metals that they form during the supernovae). The injection of energy into the surrounding gas provides a self-regulating mechanism for star formation, temporarily halting star formation until the gas radiates away the energy and collapses. The metals injected into the gas diffuse through the galaxy leading to future generations of stars having an initial metal content greater than that of previous generations.

This general trend of enrichment in galaxies with each subsequent generation of stars leads to the mass-metallicity relation (e.g. Lequeux et al., 1979; Tremonti et al., 2004; Zahid et al., 2013). Where galaxies with larger stellar mass tend to be more metal-rich than galaxies with lower stellar mass. While it is true that the average metallicity of stars increases with each subsequent enrichment event, different stars of the same age can have vastly different elemental abundances. This is due to stellar feedback being a local enrichment event in conjunction with mixing timescales in galaxies. If turbulence in a galaxy's gas is insufficient to fully mix metals before new star formation occurs then stars will have unique abundances

depending on where they formed. This process can occur in both the early and the late universe.

In the early universe, merger activity can lead to asymmetric metal distributions in galaxies, i.e. a local over- or under-abundance of metals in the region where the merger occurred. This occurs either through the dilution of metals in gas-rich mergers which subsequently trigger a star-formation event or through the accretion of stars from the merging galaxy, which will have a unique enrichment history. Additionally, in the early universe, a MW-like galaxy progenitor did not have a rotationally supported disk like the present-day MW (e.g. McCluskey et al., 2023). The lack of a rotationally supported disk contributed to the asymmetric enrichment of galaxies as gas on one side of a galaxy need not mix with gas on the opposite side.

Observations of nearby galaxies show that radial abundance gradients are well established (e.g. Belfiore et al., 2017). This is also understood in the context of theoretical models (e.g. Sharda et al., 2023) and in simulations (e.g. Bellardini et al., 2021). The presence of a radial abundance gradient necessitates that stars that form at different radii will form with different elemental abundances. Observations of the local universe also imply that the mixing timescales of galaxies are sufficient that all stars which form at a given radius will be similar in abundance (e.g. Kreckel et al., 2019, 2020).

1.3. Galactic Archaeology

Given the complexity of galaxy evolution, it should come as no surprise that researchers are using a multitude of techniques to understand the formation and evolution of galaxies. The science of using present-day observables to understand the evolutionary history of a galaxy is called galactic archaeology.

It was not even certain that galactic archaeology could be studied until the influential work of Eggen et al. (1962). They measured the kinematics and metal abundances of halo stars and determined that the metal content and orbital angular momenta of stars decreased

while at the same time, their orbital energies and eccentricities increased. Using this information, they argued for a smooth monolithic collapse of a protogalactic cloud formed the galactic stellar halo. While later work challenged this (e.g. Searle, 1977; Searle & Zinn, 1978; Marquez & Schuster, 1994) and established the current school of thought, that the formation of the halo came from a hierarchical buildup over a long period of time, the viability of galactic archaeology was established.

There are a multitude of techniques used today to try to understand galactic archaeology. Some are based on modeling the chemical evolution of galaxies (e.g. Kobayashi et al., 2020). Others involve identifying accreted substructures through precise age and stellar kinematics measurements (e.g. Koppelman et al., 2019; Massari et al., 2019; Naidu et al., 2020). Some involve leveraging precise measurements of stellar abundances in conjunction with machine learning algorithms (e.g. Ting & Weinberg, 2022). Still others involve orbit modelling (e.g. Bland-Hawthorn et al., 2019; Cantat-Gaudin et al., 2020). There is no one best method as all have associated pros and cons and can inform us about different things. The method discussed throughout this dissertation is a technique known as chemical tagging.

1.4. Chemical Tagging

The distinct elemental abundance patterns with which stars can form have initiated the development of 'chemical tagging' (Freeman & Bland-Hawthorn, 2002). This is an exceptionally powerful tool for understanding the evolutionary history of the MW as it promises to identify co-natal stars, i.e. stars that formed together, using present-day observables, e.g. their elemental abundances, kinematics, and ages. There are different applications of chemical tagging, with some researchers focusing on 'strong' chemical tagging (e.g. Hogg et al., 2016; Price-Jones et al., 2020), i.e. identifying co-natal stellar clusters. This is predicated on the assumption that star-forming clusters are essentially homogeneous (e.g. Bovy, 2016) and that star clusters form with sufficiently distinct elemental abundances such that the

elemental abundances of stars in one cluster are measurably different from those in another cluster.

There are also researchers working on a less restrictive form of chemical tagging, 'weak' chemical tagging and simply trying to associate stars with a general birth location in the disk (e.g. Anders et al., 2017) or identifying accreted stellar populations (e.g. Buder et al., 2022; Tronrud et al., 2022). This form of chemical tagging relaxes the assumption that stellar clusters must be chemically homogeneous and chemically distinct, it instead requires global abundance variations to exist in the galaxy, for in-situ chemical tagging, and unique enrichment histories/ stellar kinematics for accreted stellar populations. *Crucially, elemental abundance variations in galaxies must exist across all of cosmic time in order to chemically tag stars of any age.*

If the criteria are met, for either 'strong' or 'weak' chemical tagging one can think of the elemental abundances of a star as essentially stellar DNA. Stars that were born at the same time in the same region will have very similar, or identical DNA. Stars that are born at different times or in different regions of the galaxy will have unique DNA. Thus, with sufficient measurement precision of enough elemental abundances (for more discussion see Casamiquela et al., 2021; Ting & Weinberg, 2022; Ratcliffe et al., 2022) and stellar ages, one could in principle identify co-natal stars.

While observations of the MW and of nearby external galaxies provide us with information about the present-day state of elemental abundance variations in galaxies (e.g. Hayden et al., 2014; Belfiore et al., 2017; Wang et al., 2019; Kreckel et al., 2020), observations of the state of elemental abundance homogeneity, or lack thereof, at earlier times is harder to quantify. It is non-trivial to infer the spatial distribution of elemental abundances in the past through observations today. This is because stars experience perturbations to their orbits between their birth and present-day (e.g. Brook et al., 2004; Loebman et al., 2011; Zhang et al., 2021; Netopil et al., 2022; Lian et al., 2022, and many others). The best observational

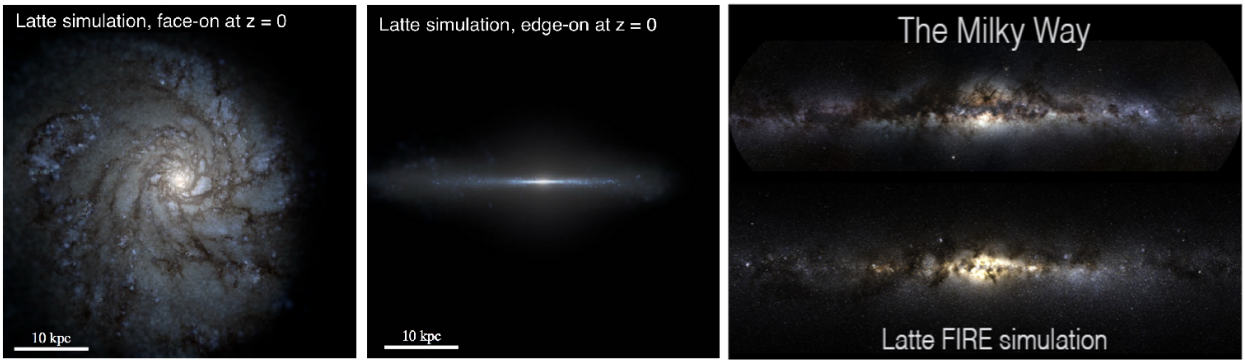


FIGURE 1.1. **Left:** Face-on mock Hubble Space Telescope image of one of the MW-like galaxies from the *Latte* suite of FIRE cosmological simulations (Wetzel et al., 2016) The spiral structure of the simulation is morphologically similar to observed spirals. **Center:** Edge-on view of the same galaxy; it exhibits a thin+thick disk morphology like the MW. **Right:** Mock, real-color, image of a Latte simulation (generated using the Ananke framework of Sanderson et al. (2020)), including an analytical prescription for dust extinction and individual stars sampled from the mass functions of the simulated star particles, compared to the MW.

techniques involve spatially resolved measurements of high redshift galaxies (e.g. Curti et al., 2020; Wang et al., 2020). However, this comes with challenges like angular resolution limits (e.g. Yuan et al., 2013).

1.5. Understanding Galaxy Evolution with Cosmological Simulations

Some of the most powerful tools for understanding the extent of elemental abundance variations across cosmic time are cosmological simulations. They provide theorists with a framework to study galactic evolution across all of cosmic time. The Feedback In Realistic Environments (FIRE) simulations are state-of-the-art hydrodynamic zoom-in simulations that utilize a mesh-free Lagrangian solver to achieve mass resolution of $\approx 7 \times 10^3 M_\odot$ (e.g. Hopkins et al., 2018). Fig. 1.1 shows an example image of these high-resolution cosmological simulations.

Crucial to characterizing the spatial distribution of elemental abundances, the FIRE simulations model sub-grid diffusion/mixing of metals in gas which occurs from unresolved

turbulent eddies (e.g. Su et al., 2017; Escala et al., 2018). This leads to spatial variations of elemental abundances in the FIRE simulations which are consistent with observations on both MW-mass scales (e.g. Bellardini et al., 2021, 2022) and dwarf galaxy scales (e.g. Porter et al., 2022). Thus, simulations can constrain the relative homogeneity of star-forming regions and newly formed stars across cosmic time. This places realistic constraints on chemical tagging models through the characterization of the evolution of physically motivated elemental abundance variations.

1.6. Dissertation Overview

In this dissertation, I present my work on the FIRE simulations, studying both the evolution of spatial distributions of elemental abundances in galaxies as well as stellar radial redistribution to better constrain chemical tagging. Chapter 2 presents the evolution of 3-D gas-phase elemental abundances in the FIRE simulations, emphasizing the steepening of radial gradients, the importance of accounting for azimuthal abundance variations in galactic chemical evolution models, and tests of the sensitivity of these results to the strength of the sub-grid diffusion coefficient. Chapter 3 extends upon the work of Chapter 2, by characterizing the 3-D spatial distribution of elemental abundances of newly formed stars across cosmic time; importantly, Chapter 3 presents functional forms and best fits to radial gradients, azimuthal scatter, and average stellar abundance as a function of time. Chapter 4 explores the radial redistribution of stars in MW-mass galaxies as a function of present-day location, formation location, and age and compares the results for 5 commonly used definitions of stellar radius.

CHAPTER 2

3-D gas-phase elemental abundances across the formation histories of Milky Way-mass galaxies in the FIRE simulations: initial conditions for chemical tagging

Published as Matthew Bellardini, Andrew Wetzel, Sarah Loebman, Claude-André Faucher-Giguère, Xiangcheng Ma, and Robert Feldmann in Monthly Notices of the Royal Astronomical Society Volume 505, Issue 3, August 2021, Pages 4586-4607, `https://doi.org/10.1093/mnras/stab1606`

2.1. Abstract

We use FIRE-2 simulations to examine 3-D variations of gas-phase elemental abundances of [O/H], [Fe/H], and [N/H] in 11 MW and M31-mass galaxies across their formation histories at $z \leq 1.5$ ($t_{\rm lookback} \leq 9.4\,{\rm Gyr}$), motivated by characterizing the initial conditions of stars for chemical tagging. Gas within 1 kpc of the disk midplane is vertically homogeneous to $\lesssim 0.008\,{\rm dex}$ at all $z \leq 1.5$. We find negative radial gradients (metallicity decreases with galactocentric radius) at all times, which steepen over time from $\approx -0.01\,{\rm dex\,kpc^{-1}}$ at $z = 1$ ($t_{\rm lookback} = 7.8\,{\rm Gyr}$) to $\approx -0.03\,{\rm dex\,kpc^{-1}}$ at $z = 0$, and which broadly agree with observations of the MW, M31, and nearby MW/M31-mass galaxies. Azimuthal variations at fixed radius are typically 0.14 dex at $z = 1$, reducing to 0.05 dex at $z = 0$. Thus, over time radial gradients become steeper while azimuthal variations become weaker (more homogeneous). As a result, azimuthal variations were larger than radial variations at $z \gtrsim 0.8$ ($t_{\rm lookback} \gtrsim 6.9\,{\rm Gyr}$). Furthermore, elemental abundances are measurably homogeneous (to $\lesssim 0.05$ dex) across a radial range of $\Delta R \approx 3.5\,{\rm kpc}$ at $z \gtrsim 1$ and $\Delta R \approx 1.7$ kpc at $z = 0$. We

also measure full distributions of elemental abundances, finding typically negatively skewed normal distributions at $z \gtrsim 1$ that evolve to typically Gaussian distributions by $z = 0$. Our results on gas abundances inform the initial conditions for stars, including the spatial and temporal scales for applying chemical tagging to understand stellar birth in the MW.

2.2. Introduction

Many current and future observational surveys of stars across the Milky Way (MW) seek to unveil the MW's formation history in exquisite detail. Current surveys, such as the RAdial Velocity Experiment (RAVE; Steinmetz et al., 2006), the Gaia-ESO survey (Gilmore et al., 2012), the Large Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Cui et al., 2012), GALactic Archaeology with Hermes (GALAH; De Silva et al., 2015), and the Apache Point Galactic Evolution Experiment (APOGEE; Majewski et al., 2017) have measured elemental abundances of hundreds of thousands of stars. Future surveys, such as the WHT Enhanced Area Velocity Explorer (WEAVE; Dalton et al., 2012), the Subaru Prime Focus Spectrograph (PFS; Takada et al., 2014), the Sloan Digital Sky Survey V (SDSS-V; Kollmeier et al., 2017), the 4-metre Multi-Object Spectrograph Telescope (4MOST; De Jong et al., 2019), and the MaunaKea Spectroscopic Explorer (MSE; The MSE Science Team et al., 2019) will increase the samples of spectroscopically observed stars into the millions. A key science driver for these surveys is 'galactic archaeology': to infer the history of the MW using observations of the dynamics and elemental abundances of stars today.

Measurements of stellar dynamics can provide detailed information on the MW's properties and formation history, but the fundamental limitation is that stellar orbits can change over time via mergers, accretion, scattering, and other dynamical perturbations (e.g. Abadi et al., 2003; Brook et al., 2004; Schönrich & Binney, 2009; Loebman et al., 2011). However, a star's atmospheric elemental abundances will not change in response to these dynamical

processes, providing a key orbital-invariant 'tag'. 'Chemical tagging', introduced in Freeman & Bland-Hawthorn (2002), thus provides tremendous potential to infer the formation conditions of a star of arbitrary age.

'Strong' chemical tagging represents the most fine-grained scenario, to identify stars born in the same star cluster (e.g. Price-Jones et al., 2020). By contrast, 'weak' chemical tagging seeks to infer the general location and time where/when a stellar population formed, for example, to associate populations of stars to certain birth regions of the galaxy (e.g. Wojno et al., 2016; Anders et al., 2017) or that accreted into the MW from galaxy mergers (e.g. Ostdiek et al., 2020).

Both regimes of chemical tagging rely on sufficiently precise measurements of stellar abundances and on assumptions about the elemental homogeneity (to the measured precision) of the gas from which the stars formed. For example, strong chemical tagging of individual star clusters relies on both the internal homogeneity of the gas cloud from which the stars formed, and on how unique the abundance patterns were in that cloud across space and time. Observational evidence of open star clusters suggests the first criterion is met (De Silva et al., 2007; Ting et al., 2012; Bovy, 2016) to measurable precision. Regarding the latter criterion, observations of the MW and external galaxies show radial and azimuthal variations in abundances across the disk (e.g. Sánchez-Menguiano et al., 2016; Mollá et al., 2019b; Wenger et al., 2019; Kreckel et al., 2020), although more work is needed to understand these spatial variations in the context of chemical tagging. Weak chemical tagging is subject to the same assumptions but applied to larger regions of gas across the disk (or in accreting galaxies). For example, if all gas in the disk was measurably homogeneous in all abundances at a given time, chemical tagging would offer no spatially discriminating power. Conversely, the limit of extreme clumpiness, in which each star cluster formed with a measurably unique abundance pattern, would in principle enabled detailed chemical tagging, but it significantly would complicate the modeling.

Thus, a key question for chemical tagging is: what are the relevant spatial scales of measurable homogeneity of stars forming at a given time, and how does this evolve across cosmic time?. Bland-Hawthorn et al. (2010) previously explored this via a toy model, where they show all star clusters $\lesssim 10^4 M_\odot$ and a large fraction of clusters with mass below $\sim 10^5 M_\odot$ are expected to be internally homogeneous. Progress in chemical tagging requires addressing these questions regarding stellar birth before examining the subsequent dynamical evolution of stars after they form.

Many works have examined abundance variations of stars across the MW, generally finding a negative radial gradient in abundances for stars near the plane of the disk, which flattens or turns positive at larger heights (e.g. Cheng et al., 2012; Boeche et al., 2013, 2014; Anders et al., 2014; Hayden et al., 2014; Mikolaitis et al., 2014; Anders et al., 2017). Furthermore, the MW has an observed negative vertical gradient in stars (Cheng et al., 2012; Carrell et al., 2012; Boeche et al., 2014; Hayden et al., 2014), although this slope varies significantly between observations and depends on radius (Hayden et al., 2014). Both the radial and vertical gradients vary with stellar age (Wang et al., 2019). Additionally, Luck et al. (2006); Lemasle et al. (2008); Pedicelli et al. (2009) found evidence for azimuthal variations in the abundances of young stars, which may result from patchy star formation (Davies et al., 2009; Luck & Lambert, 2011; Genovali et al., 2014). Nieva & Przybilla (2012) also explored the homogeneity of B-type stars in the solar neighborhood, within 500 pc, of the sun and found scatter on the order of 0.05 dex for [O/H] which they state is comparable to gas-phase abundance scatter out to 1.5 kpc of the sun.

In addition to stellar abundances, many works have characterized trends of gas-phase abundances in the MW. Observations show that the MW has a negative radial gradient in gas-phase abundances, with a slope that varies across the elements (Arellano-Córdova et al., 2020) and across studies (e.g. Mollá et al., 2019a, and references therein). Furthermore,

evidence persists for azimuthal variations in this radial gradient, based on HII regions (Balser et al., 2011, 2015; Wenger et al., 2019).

Beyond the MW, observations of nearby MW-mass galaxies also show negative radial gradients in gas-phase abundances (e.g. Pilyugin et al., 2014; Sánchez-Menguiano et al., 2016; Belfiore et al., 2017; Poetrodjojo et al., 2018). Furthermore, some observations show azimuthal variations (Sánchez et al., 2015; Vogt et al., 2017; Ho et al., 2017, 2018; Kreckel et al., 2019, 2020; Sánchez-Menguiano et al., 2020), while others show no azimuthal variations within measurement uncertainty ($\lesssim 0.05\,\mathrm{dex}$) (e.g. Cedrés & Cepa, 2002; Zinchenko et al., 2016).

Understanding how these variations change across cosmic time is imperative for chemical-tagging models. Currently no consensus exists, amongst observations (e.g. Curti et al., 2020, and references therein) on the redshift evolution of radial elemental abundance gradients, in part because of angular resolution limitations (Yuan et al., 2013), which some works have addressed via adaptive optics and gravitational lensing (Jones et al., 2010; Swinbank et al., 2012; Jones et al., 2013). Furthermore, different works use different calibrators to measure abundances, which often disagree (Hemler et al., 2021), further complicating our understanding of spatial variations.

Many theoretical works have used simulations to predict the spatial distribution of gas-phase abundances and their evolution. As with observational efforts, there is no consensus for the redshift evolution of abundance gradients in theory (e.g. Mollá et al., 2019a, and references therein). Gibson et al. (2013) compared cosmological simulations with MUGS ('conservative' feedback) and MaGICC ('enhanced' feedback) run with the GASOLINE code and found that the strength of feedback in simulations is critical for the evolution of radial gradients of abundances, such that stronger feedback leads to flatter gradients at all times, while galaxies with weaker feedback have gradients that are steep at high redshift and flatten with time. Ma et al. (2017a), studying the FIRE-1 suite of cosmological simulations,

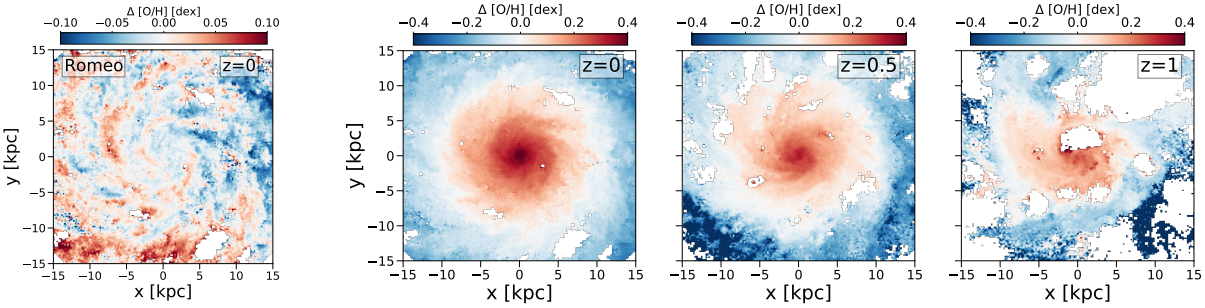


FIGURE 2.1. Face-on image of all gas within ±1 kpc of the galactic midplane of Romeo, one of FIRE-2 simulations that we analyze. We color-code gas by Δ[O/H], its deviation from the azimuthally averaged [O/H]. Δ[Fe/H] (not shown) looks nearly identical to Δ[O/H], to within $\lesssim 0.02$ dex at all times. The left panel emphasizes azimuthal variations by showing the deviation from the azimuthally averaged [O/H] *at each radius*, that is, subtracting off the radial gradient, thus highlighting the enhanced enrichment correlated with spiral arms. The right 3 panels show the deviation from the mean [O/H] of all gas at $R \leq 15$ kpc at each redshift. White regions have highly diffuse gas in which we do not report a measured abundance. The radial gradient in [O/H] dominates over azimuthal variations at late times, but at early times the azimuthal variations are the most significant.

found that galaxies exhibit a diverse range of radial gradients in abundances, and that these gradients can fluctuate rapidly from steep to shallow (in ~ 100 Myr) at high redshift, so measurements of high-redshift gradients may not be indicative of long-term trends. They found that galaxies tend to quickly build up a negative gradient once stellar feedback is no longer sufficient to drive strong outflows of gas. By contrast, analyzing star-forming galaxies in the TNG-50 cosmological simulation, Hemler et al. (2021) found that radial gradients in galaxies are steep at high redshift and flatten with time. Several theoretical works also have examined azimuthal variations. Spitoni et al. (2019) developed a 2-D model for abundance evolution that follows radial and azimuthal density variations in a MW-like disk and found that azimuthal residuals are strongest at early times and at large radii. Using their S2A model, Spitoni et al. (2019) found azimuthal residuals in [O/H] of ≈ 0.1 dex at $R = 8$ kpc at $t_{\rm lookback} = 11$ Gyr which evolve to ≈ 0.05 dex at present day. Mollá et al. (2019b) explored azimuthal variations in a MW-like disk for 5 models of 2-D abundance evolution and found

[O/H] variations that are typically small ($0.05 - 0.1$ dex) and dilute quickly with time. Solar et al. (2020) used young star particles as tracers of star-forming gas in the EAGLE cosmological simulation and found an average azimuthal abundance dispersion of ≈ 0.12 dex at $z = 0$ in galaxies with $M_* = 10^9 - 10^{10.8} \mathrm{M}_\odot$.

Observations of azimuthal variations of abundances in gas in nearby MW-mass galaxies find scatter that is comparable to observational measurement uncertainty (~ 0.05 dex) (Zinchenko et al., 2016; Kreckel et al., 2019). This implies that gas in galaxies is well mixed azimuthally at $z = 0$. Consequently, works modeling the abundance evolution of galaxies, which inform chemical tagging, generally assume that gas is well mixed azimuthally in the disk at all times (e.g. Minchev et al., 2018; Mollá et al., 2019a; Frankel et al., 2020), such that the key spatial variation is radial. While galaxies exhibit radial gradients across a range of redshifts (Queyrel et al., 2012; Stott et al., 2014; Wuyts et al., 2016; Carton et al., 2018; Patrício et al., 2019; Curti et al., 2020), radial variations may not always dominate. At early times, in particular, azimuthal variations may be more important. Some abundance-evolution models have begun to explore both azimuthal and radial variations across time (e.g. Acharova et al., 2013; Mollá et al., 2019b; Spitoni et al., 2019). Kawata et al. (2014); Grand et al. (2015), using N-body simulations of MW-mass galaxies, and Baba et al. (2016), using baryonic simulations run with the SPH code ASURA-2, found that gas exhibits motion along spiral arms, which could contribute to 2-D abundance variations in gas.

More detailed 2-D abundance-evolution models, which account for density variations within the disk from spiral arms and bars, result in azimuthal variations in gas-phase abundances. Mollá et al. (2019b) found that arm / inter-arm abundance variations quickly dilute through interactions with spiral structure. Spitoni et al. (2019) also found that azimuthal variations dilute with time, but they found that the strength of azimuthal variations at $z = 0$ approximately agree with the observational results of Kreckel et al. (2019). The simulation analysis of Grand et al. (2015); Baba et al. (2016) showed that, just like stars (Lynden-Bell

& Kalnajs, 1972), gas experiences radial migration as a result of spiral structure. Grand et al. (2015) found that this systematic streaming along spiral arms leads to metal-rich gas in the inner galaxy moving to larger radii and metal-poor gas in the outer galaxy moving to the inner galaxy, leading to non-homogeneous abundances at a given radius. However, they found that gas elements quickly exchange abundances after migrating, leading to small azimuthal dispersions in abundance. Sánchez-Menguiano et al. (2020) found that azimuthal variations in abundances are stronger in galaxies with stronger bars and grand-design spirals, which supports non-axisymmetric structure driving azimuthal inhomogeneities.

In this paper we use FIRE-2 cosmological simulations of MW/M31-mass galaxies to explore the cosmic evolution of 3-D abundance patterns of gas, as a first step towards understanding the spatial and temporal scales of applying chemical tagging in a cosmological context. Our analysis of gas represents our first step, to characterize the initial conditions for star-forming regions. In future work, we will examine the resultant trends in stars and their dynamical evolution across time. Here, we seek to quantify the 3-D spatial scales over which elemental abundances of gas (and thus the formation of stars) are measurably homogeneous. In Section. 2.3 we describe the simulations used for this analysis. In Section. 2.4 we first explore the radial gradients and compare them against observations of the MW, M31, and nearby MW/M31-mass galaxies. Next we examine the cosmic evolution of radial, vertical, and azimuthal variations in gas-phase abundances, in particular, to understand which dimension dominates the spatial variations at a given time. We also examine implications of gas (in)homogeneity on current and upcoming observations of the MW. Finally, we examine full distributions of elemental abundances. Section. 2.5 we summarize the main results of the paper and provide a discussion of their implications.

2.3. Methods

simulation	M_{90}^{star} [10^{10} $M_\odot$]	R_{25}^{star} [kpc]	R_{50}^{star} [kpc]	R_{90}^{star} [kpc]	Z_{90}^{star} [kpc]	M_{90}^{gas} [10^{10} $M_\odot$]	R_{25}^{gas} [kpc]	R_{50}^{gas} [kpc]	R_{90}^{gas} [kpc]	Z_{90}^{gas} [kpc]	f_{90}^{gas} ($< R_{90}^{star}$, $< Z_{90}^{star}$)
m12m	10.0	1.9	4.3	11.6	2.3	2.1	6.6	10.3	15.0	1.2	0.13
Romulus	8.0	1.2	3.2	12.9	2.4	2.7	9.0	13.1	18.3	2.3	0.16
m12b	7.3	1.0	2.2	9.0	1.8	1.7	6.4	9.6	15.0	1.5	0.11
m12f	6.9	1.2	2.9	11.8	2.1	2.3	8.7	12.6	17.8	2.4	0.14
Thelma	6.3	6.3	3.4	11.2	3.2	2.6	7.5	12.1	17.6	3.1	0.16
Romeo	5.9	1.6	3.6	12.4	1.9	1.8	8.0	12.2	18.1	1.5	0.16
m12i	5.3	1.1	2.6	9.8	2.3	1.7	7.1	10.2	16.7	1.7	0.15
m12c	5.1	1.3	2.9	9.1	2.0	1.5	5.4	8.4	14.6	2.4	0.15
Remus	4.0	1.2	2.9	11.0	2.2	1.5	7.4	11.8	18.0	1.8	0.22
Juliet	3.3	0.8	1.8	8.1	2.2	1.5	6.9	11.3	18.6	3.1	0.14
Louise	2.3	1.2	2.8	11.2	2.2	1.4	8.0	12.6	18.5	2.0	0.34

TABLE 2.1. Properties of the stellar and gas disks of our simulated MW/M31-mass galaxies $z = 0$. The first column lists the name of the galaxy: 'm12' indicates isolated galaxies from the *Latte* suite, while the other galaxies are LG analogues from the ELVIS on FIRE suite. R_{25} and R_{50} is the radius where the cumulative mass of the disk is 25% and 50%, respectively, within a height ± 3 kpc out of the midplane, relative to the total stellar/gas mass of the disk within 20 kpc. We fit R_{90} and Z_{90} as the radius and height where the cumulative mass of the stellar/gas disk are 90% of the total mass of stars/gas within a sphere of 20 kpc. M_{90} is the total stellar/gas mass contained within both R_{90} and Z_{90}. The gas fraction, f_{90}^{gas}, is the ratio of gas mass to total baryonic mass within R_{90}^{star} and Z_{90}^{star}.

2.3.1. FIRE-2 Simulations. We use a suite of MW/M31-mass cosmological zoom-in simulations from the Feedback In Realistic Environments (FIRE) project[1] (Hopkins et al., 2018). We ran these simulations using the FIRE-2 numerical implementations of fluid dynamics, star formation, and stellar feedback. These simulations use the Lagrangian Meshless Finite Mass hydrodynamics method in GIZMO (Hopkins, 2015). The FIRE-2 model incorporates physically motivated metallicity-dependent radiative heating and cooling processes for gas such as free-free, photoionization and recombination, Compton, photo-electric and dust collisional, cosmic ray, molecular, metal-line, and fine structure processes, accounting for 11 elements (H, He, C, N, O, Ne, Mg, Si, S, Ca, Fe) across a temperature range of $10 - 10^{10}$K. The simulations also include a spatially uniform, redshift-dependent UV background from Faucher-Giguère et al. (2009). In calculating metallicities throughout this paper, we scale elemental abundances to the solar values in Asplund et al. (2009).

Star particles form out of gas that is self-gravitating, Jeans-unstable, cold ($T < 10^4$K), dense ($n > 1000\,\mathrm{cm}^{-3}$), and molecular (following Krumholz & Gnedin, 2011). Each star particle inherits the mass and elemental abundances of its progenitor gas and represents a single stellar population, assuming a Kroupa (2001) stellar initial mass function. FIRE-2 evolves star particles along standard stellar population models from e.g. STARBURST99 v7.0 (Leitherer et al., 1999), including time-resolved stellar feedback from core-collapse and Ia supernovae, continuous mass loss, radiation pressure, photoionization, and photo-electric heating. FIRE-2 uses rates of core-collapse and Ia supernovae from STARBURST99 (Leitherer et al., 1999) and Mannucci et al. (2006), respectively. The nucleosynthetic yields follow Nomoto et al. (2006) for core-collapse and Iwamoto et al. (1999) for Ia supernovae. Stellar wind yields, sourced primarily from O, B, and AGB stars, are from the combination of models from Van den Hoek & Groenewegen (1997); Marigo (2001); Izzard et al. (2004), synthesized in Wiersma et al. (2009).

[1]FIRE project web site: http://fire.northwestern.edu

Critical for this work, these FIRE-2 simulations also explicitly model the sub-grid diffusion/mixing of elemental abundances in gas that occurs via unresolved turbulent eddies (Su et al., 2017; Escala et al., 2018; Hopkins et al., 2018). In effect, this smooths abundance variations between gas elements, assuming that sub-grid mixing is dominated by the largest unresolved eddies. Escala et al. (2018) showed that incorporating this sub-grid model is a necessity to match observed distributions of stellar metallicities. We explore the robustness of our results to variations in the strength of the mixing/diffusion coefficient in Appendix A.3.

All simulations assume flat ΛCDM cosmologies with parameters broadly consistent with the Planck Collaboration et al. (2020): $h = 0.68-0.71$, $\Omega_\Lambda = 0.69-0.734$, $\Omega_\mathrm{m} = 0.266-0.31$, $\Omega_\mathrm{b} = 0.0455 - 0.048$, $\sigma_8 = 0.801 - 0.82$ and $n_s = 0.961 - 0.97$. For each simulation we generated cosmological zoom-in initial conditions embedded within cosmological boxes of length $70.4 - 172\,\mathrm{Mpc}$ at $z \approx 99$ using the code MUSIC (Hahn & Abel, 2011). We saved 600 snapshots from $z = 99$ to 0, with typical time spacing $\lesssim 25\,\mathrm{Myr}$.

We examine 11 MW/M31-mass galaxies from 2 suites of simulations. We select only galaxies with a stellar mass within a factor of ≈ 2 of the MW, $\approx 5 \times 10^{10}\mathrm{M}_\odot$ (Bland-Hawthorn & Gerhard, 2016). 5 of our galaxies are from the *Latte* suite of isolated individual MW/M31-mass halos, introduced in Wetzel et al. (2016). (We exclude m12w, because it has an unusually compact gas disk at $z = 0$, with $R_{90}^{\mathrm{gas}} = 7.4\,\mathrm{kpc}$). *Latte* galaxies have halo masses $M_{200\mathrm{m}} = 1 - 2 \times 10^{12}\mathrm{M}_\odot$, for which $M_{200\mathrm{m}}$ refers to the total mass within the radius containing 200 times the mean matter density of the Universe. These simulations have Dark Matter (DM) particle masses of $3.5 \times 10^4\mathrm{M}_\odot$ and initial baryonic particle masses of $7070\mathrm{M}_\odot$ (though because of stellar mass loss, star particles at $z = 0$ typically have masses of $\sim 5000\mathrm{M}_\odot$). Star and DM particles have fixed gravitational force softening lengths of 4 and 40 pc (Plummer equivalent), comoving at $z > 9$ and physical thereafter. Gas elements have adaptive gas smoothing and gravitational force softening lengths that reach a minimum of 1 pc. We also include 6 galaxies from the 'ELVIS on FIRE' suite (Garrison-Kimmel et al.,

2014, 2019a). These simulate Local Group (LG)-like MW+M31 pairs. ELVIS hosts have halo masses $M_{200\mathrm{m}} = 1-3\times10^{12}\mathrm{M}_{\odot}$, with $\approx 2\times$ better mass resolution than the *Latte* suite.

In general, we find few systematic differences in any of our results for the isolated galaxies versus the LG-like pairs, the only notable difference being the relative strength of azimuthal scatter to radial gradient strength at large radius and high z, so we combine these suites in all of our results.

2.4. Results

Fig. 2.1 shows face-on images of the gas disk of one of our simulations, Romeo, at several redshifts. We color-code gas by its variation in [O/H], to visualize key trends that we explore in this work. We do not show results for [Fe/H], because they are qualitatively consistent with [O/H]. The left panel shows the deviation of the local [O/H] from the mean value *at each radius* for radial bins of width 200 pc at $z = 0$, that is, we subtract off the overall radial gradient. This highlights the variations along the azimuthal direction at each radius, showing enhancement in [O/H] along spiral structure (see Orr et al., 2023).

The right 3 panels show the difference between the local [O/H] and the mean [O/H] across all gas at $R \leq 15$ kpc and within ±1 kpc of the galactic midplane at each redshift. This highlights the importance of both radial and azimuthal abundance variations in gas. At late times, the gas disk shows a clear negative radial gradient that is much stronger than the azimuthal variations. However, at earlier times, the gas disk is azimuthally more asymmetric, including cavities from local star-forming and feedback regions. A radial gradient is less pronounced. As we will show, at $z \gtrsim 0.8$ ($t_{\mathrm{lookback}} \gtrsim 6.9$ Gyr) the azimuthal variations in abundance at a given radius are typically larger than the radial change across the disk.

Because the absolute normalization of any elemental abundance in our simulations is uncertain, given uncertainties in underlying nucleosynthetic rate and yield models, throughout this paper we focus on *relative* abundance variations, including spatial variations, evolution,

and the shapes of abundance distributions rather than absolute normalizations of abundances.

As Fig. 2.2 shows, gas [O/H] in our simulations is super-solar at all radii (out to 15 kpc). Over comparable radial ranges, our hosts have an average gas-phase [O/H] ≈ 0.56 dex larger than what is observed in the MW (Fernández-Martín et al., 2017; Wenger et al., 2019; Arellano-Córdova et al., 2020) and M31 (Zurita & Bresolin, 2012; Sanders et al., 2012). Most likely the primary cause of this high normalization is our modeling of core-collapse supernovae in FIRE-2. We assume all core-collapse supernovae to have identical (IMF-averaged) yields, but in reality, different mass progenitors produce different yields, as Muley et al. (2021) explored in the context of FIRE-2.

Furthermore, in FIRE-2 we assume nucleosynthetic yields for core-collapse supernovae from Nomoto et al. (2006), but more recent compilations (Nomoto et al., 2013; Pignatari et al., 2016; Sukhbold et al., 2016; Limongi & Chieffi, 2018) suggest $\sim 2\times$ lower O yields (IMF-averaged), as we will show in detail with our next-generation FIRE model (Hopkins et al., 2023). In that work, we also will show that the overall stellar-wind mass-loss rates are likely $\sim 2\times$ smaller than in FIRE-2, which further contributes to lower [O/H]. Finally, as we will explore in Gandhi et al. (2022), our assumed supernovae Ia rates in FIRE-2 (Mannucci et al., 2006) are $\sim 2\times$ lower than more recent Ia rate constraints (e.g. Maoz & Graur, 2017). Thus, these updates will account for up to $\sim 0.3-0.4$ dex lower normalization in our predicted [O/H], and up to $0.4-0.5$ dex lower normalization in our [O/Fe].

2.4.1. Radial profiles at z = 0. First we examine the radial profiles of [O/H], [Fe/H], and [O/Fe] in gas for all 11 galaxies at $z = 0$. We time-average each galaxy's profile across ~ 50 Myr by stacking 3 snapshots to reduce short-time fluctuations. We present all results in physical radii; in Appendix A.1, we examine these trends scaling to various galactic scale radii, finding that the host-to-host scatter in our suite is minimized when examining gradients in physical units. These profiles contain all gas within a vertical height $Z \pm 1$ kpc

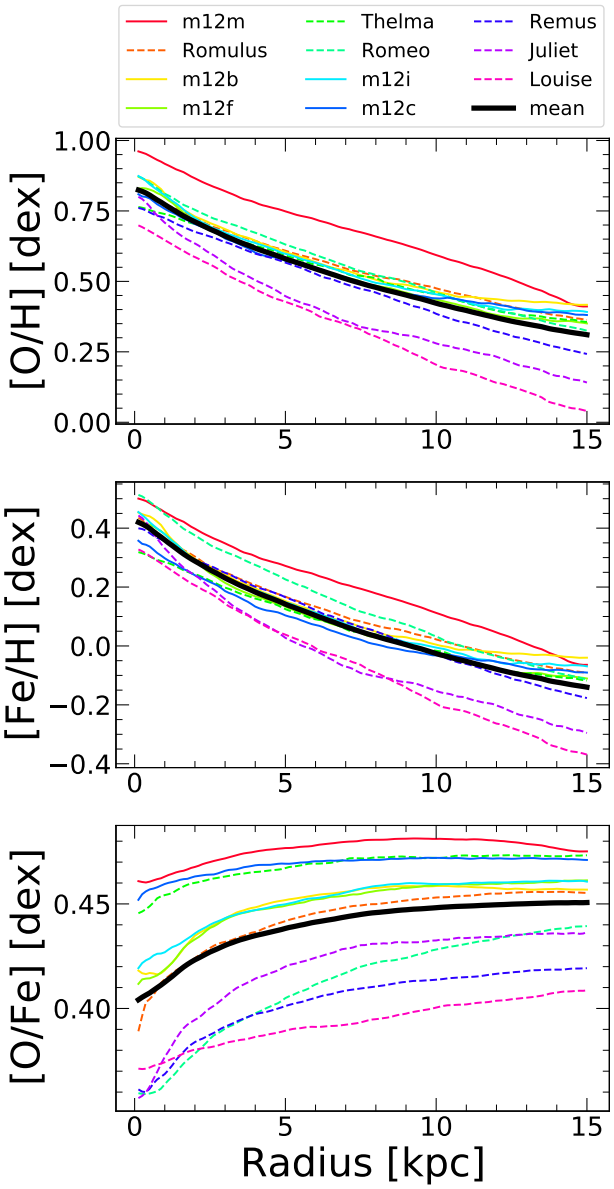


FIGURE 2.2. Radial profiles of gas-phase elemental abundances within the disks of our 11 galaxies at $z = 0$, listed by decreasing stellar mass. Each simulation includes all gas, averaging across 3 snapshots (≈ 50 Myr) within a disk height of ± 1 kpc. The black line shows the mean across all hosts. The normalization of the radial abundances scales with the stellar mass of the galaxy (see Table: 2.1). For [O/H] and [Fe/H], the profiles exhibit negative radial gradients, with a mean change across $0 - 15$ kpc of ~ 0.51 dex for [O/H] and ~ 0.56 dex for [Fe/H]. The [O/Fe] profile is approximately flat at radii $\gtrsim 4$ kpc, indicating enrichment dominated by core-collapse supernovae, but it exhibits a small positive gradient in the inner galaxies, from the increasing importance of Fe from Ia supernovae.

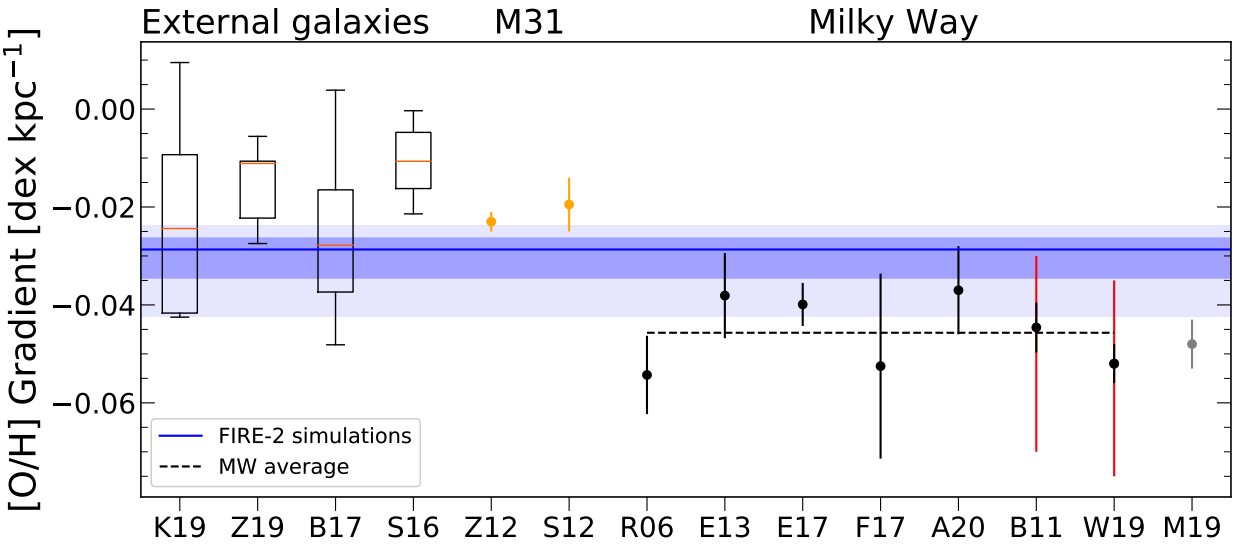


FIGURE 2.3. Radial gradients in gas-phase [O/H] across our 11 galaxies and observed in the MW, M31, and in nearby MW-mass galaxies. The blue horizontal line shows the median across our 11 galaxies, with the dark shaded region showing the 68th percentile and the light shaded region the full distribution. We also show observations of radial gradients in external galaxies, from Kreckel et al. (2019, K19), Zinchenko et al. (2019, Z19), Belfiore et al. (2017, B17), and Sánchez-Menguiano et al. (2016, S16), via box-and-whisker, where the box displays the 68th percentile, the whiskers display the full distribution, and the orange horizontal line is the median. Orange circles show observed abundance gradients for M31 derived from HII regions by Zurita & Bresolin (2012, Z12) and Sanders et al. (2012, S12). Black circles show observed abundance gradients for the MW derived from HII regions from Rudolph et al. (2006, R06), Balser et al. (2011, B11), Esteban et al. (2013, E13), Esteban et al. (2017, E17), Fernández-Martín et al. (2017, F17), Wenger et al. (2019, W19), and Arellano-Córdova et al. (2020, A20). Mollá et al. (2019a, M19), shown in grey, is the gradient derived from a compilation of the data from R06, B11, E13, E17, and F17. We show uncertainties for all points. For W19 and B11 the red shows the variation in gradient observed by looking along different azimuths. The dashed line shows the best-fit MW gradient ($-0.046\,\mathrm{dex\,kpc^{-1}}$), based on the gradients the observations presented here (excluding M19). The median [O/H] gradient across our galaxies is $-0.028\,\mathrm{dex\,kpc^{-1}}$ and the standard deviation is $0.005\,\mathrm{dex\,kpc^{-1}}$, in agreement with K19 and B17 to within 1σ, and with Z19 to within 2σ, but not in agreement with S16. Our simulations also agree with observations of M31 and some of the MW observations.

from the disk and we use radial bins of width 0.25 kpc. We calculate the mass-weighted mean of the gas-phase abundance in each bin. We show profiles out to $R = 15$ kpc; our gas disks generally extend beyond this radius, but because our primary motivation is chemical tagging, we examine only regions with significant star formation (see Table 2.1).

Fig. 2.2 (top 2 panels) shows that [O/H] and [Fe/H] decrease monotonically with radius. The mean gradient is $\approx -0.03\,\mathrm{dex\,kpc^{-1}}$ for both [O/H] and [Fe/H], and the mean change in abundance from $0 - 15$ kpc is ≈ 0.51 dex for [O/H] and ≈ 0.56 dex for [Fe/H]. These negative gradients in gas reflect the decreasing ratio of stars (the sources of enrichment) to gas towards the outer disk, and show that these gas disks are not radially well mixed at $z = 0$.

Across our 11 galaxies, the host-to-host standard deviation is ≈ 0.09 dex for [O/H] and ≈ 0.07 dex for [Fe/H]. The legend of Fig. 2.2 lists the host galaxies in decreasing order of stellar mass, highlighting that the abundance at a given radius correlates strongly with the galaxy's mass. Table. 2.1 shows that stellar mass drops by a factor of ~ 4 from m12m to Louise; given the slope of the gas-phase mass-metallicity relation from Ma et al. (2016), ≈ 0.4 dex, the scatter in [O/H] normalization for our mass range should be ≈ 0.24 dex, almost exactly that in Fig. 2.2. In other words, the scatter across our suite primarily reflects the mass-metallicity relation (see Lequeux et al., 1979; Tremonti et al., 2004; Mannucci et al., 2010; Andrews & Martini, 2013).

In Fig. 2.2, dashed lines show the LG-like hosts, and while they show typically lower abundance at a fixed radius than the isolated hosts, this is because they have somewhat lower stellar mass on average. We find no systematic differences between LG-like and isolated hosts beyond this, despite the fact that the LG-like hosts form their in-situ stars systematically earlier than the isolated hosts (Santistevan et al., 2020). Thus, this difference in formation history does not imprint itself on gas-phase abundances at $z \leq 1.5$. As a result, we will combine these samples in all subsequent analyses.

Fig. 2.2 (bottom panel) shows profiles for [O/Fe], which are nearly flat at all radii. The mean change in [O/Fe] from $0 - 15\,\mathrm{kpc}$ is $\approx -0.046\,\mathrm{dex}$. [O/Fe] shows the strongest (positive) gradient in the inner $\approx 4\,\mathrm{kpc}$, highlighting the increasing importance of enrichment from (more delayed) Ia supernovae towards the galactic center, which underwent the longest period of enrichment. However, the outer disk, beyond $\approx 4\,\mathrm{kpc}$, reflects relatively similar enrichment from core-collapse and Ia supernovae at each radius. We find a host-to-host standard deviation of $\approx 0.027\,\mathrm{dex}$ for [O/Fe]. We measure this for [Mg/Fe] (not shown here) as another tracer of core-collapse vs Ia supernovae enrichment. The mean change in [Mg/Fe] from $0 - 15\,\mathrm{kpc}$ is $\approx -.111\,\mathrm{dex}$ and the host-to-host standard deviation is $\approx 0.02\,\mathrm{dex}$. These differences are likely attributable to our stellar wind model having a metallicity dependent yield for O and not for Mg. This leads to more O production at small radii where the metallicity is higher, thus a flatter profile.

Fig. 2.3 shows the radial gradients of [O/H] in our simulations at $z = 0$ and includes observations of the MW, M31, and nearby MW/M31-mass galaxies. We fit the gradients in our simulations using a least-squares fit of the [O/H] abundance across $4 - 12\,\mathrm{kpc}$. As Fig. 2.2 shows, including the inner region of our disks, where the bulge dominates ($R \lesssim 4\,\mathrm{kpc}$), gives a profile not well approximated by a single linear fit (the bulge is steeper), so we exclude it in fitting this profile, to measure the 'disk' component. The range $4 - 12\,\mathrm{kpc}$ covers the inner and outer disk and generally exhibits a single power-law profile. The solid blue line shows the median ($-0.028\,\mathrm{dex\,kpc^{-1}}$) across our 11 galaxies, while the shaded regions show the 68th percentile and the full distribution. The latter ranges from -0.042 to $-0.024\,\mathrm{dex\,kpc^{-1}}$. [Fe/H] gradients show similar results, with the full distribution spanning -0.044 to $-0.024\,\mathrm{dex\,kpc^{-1}}$.

Fig. 2.3 shows [O/H] gradients observed in nearby MW/M31-mass galaxies as box-and-whisker plots, with the box showing the 68th percentile and the whiskers showing the full observed range. We apply a cut on the stellar masses of these observed samples to be

comparable to our simulations. The Kreckel et al. (2019, K19) sample includes 5 galaxies from the PHANGS-MUSE survey with $10.2 \leq \log_{10} M_{\rm star}/M_{\odot} \leq 10.6$, the Zinchenko et al. (2019, Z19) sample includes 7 galaxies from CALIFA DR3 with $10.2 \leq \log_{10} M_{\rm star}/M_{\odot} \leq 10.8$, the Belfiore et al. (2017, B17) sample includes 13 galaxies from the MaNGA survey with $10.2 \leq \log_{10} M_{\rm star}/M_{\odot} \leq 11$, and the Sánchez-Menguiano et al. (2016, S16) sample includes 20 galaxies from the CALIFA survey with $10.2 \leq \log_{10} M_{\rm star}/M_{\odot} \leq 11$. In addition to the mass cut, we select galaxies that have gradients measured across a radial range comparable to the range in our analysis (the measured ranges all fall within $2-14\,$kpc except for Kreckel et al. (2019) which falls within $1-11\,$kpc). While all of these observed samples show almost exclusively negative gradients in [O/H], their abundance gradients are typically flatter than in our simulations. Our simulations are consistent at the 1-σ level with K19 and B17, and at the 2-σ level with Z19. However, our sample does not overlap with S16. Note that the calibrator used for determining the abundances varies from survey to survey. Using different calibrators can give drastically different abundance measurements (Hemler et al., 2021), which could contribute to discrepancies between the different surveys, and to differences with our simulations. Note that the difference between our simulations and these observations are comparable to the differences between surveys themselves.

Fig. 2.3 also shows observed abundance gradients in M31 and the MW from HII regions. The orange points show observed gradients in M31 from Zurita & Bresolin (2012, Z12) and Sanders et al. (2012, S12). These gradients are slightly shallower than in our simulations, though they agree within 2-σ. This may be a consequence of M31 gradient measurements spanning $\approx 4-25\,$kpc: from our analysis, including the outer regions of a gas disk flattens the inferred gradient.

The black points show measured gradients of the MW. Mollá et al. (2019a), shown in grey, is a best-fit measurement of the MW abundance gradient based on the combined data of Rudolph et al. (2006); Balser et al. (2011); Esteban et al. (2013, 2017); Fernández-Martín

et al. (2017). We show uncertainties for all samples. The red error bars for Balser et al. (2011); Wenger et al. (2019) show the impact of measuring the radial gradient along different galactic azimuths. Balser et al. (2011) finds gradients ranging from -0.03 to $-0.07\,\mathrm{dex\,kpc^{-1}}$ and Wenger et al. (2019) find gradients ranging from -0.035 to $-0.075\,\mathrm{dex\,kpc^{-1}}$ which highlights that measurements of the MW radial gradient are strongly sensitive to azimuthal variations. The different samples include different radial ranges, so they are not exactly comparable to each other or our analysis. Most measurements of [O/H] gradients in the MW overlap with our simulations, though our simulations generally have shallower gradients.

While not included in Fig. 2.3, Hernandez et al. (2021) recently measured the radial [O/H] gradient in neutral and ionized gas in M83. The gradients were measured out to $\approx 5.5\,\mathrm{kpc}$. They found the gradients in neutral gas to be substantially steeper than the gradients in ionized gas. As most observations target ionized gas around HII regions, one might expect that our measured gradients shown in Fig. 2.3 are flatter than expected. However, Hernandez et al. (2021) measured gradients primarily in the bulge, which we exclude in this analysis. Their gradient for neutral gas external to the bulge is $\approx -0.02\,\mathrm{dex\,kpc^{-1}}$ and for ionized gas is $\approx -0.03\,\mathrm{dex\,kpc^{-1}}$, in good agreement with our values.

As a whole, the radial gradients in our simulations are somewhat steeper than in external galaxies but somewhat shallower than in the MW. The MW may be an outlier: as Boardman et al. (2020) note, its gradient is typically steeper than those observed in MW analogs. These differences are likely the result of a combination of different factors, such as: measuring over different radial ranges or using different calibrators. For example, B17 also measure the gradient in their MaNGA observations using a different calibrator for [O/H] (O3N2, not shown here, as opposed to R23, as Fig. 2.3 shows), which results in a median gradient that is $\approx 0.008\,\mathrm{dex\,kpc^{-1}}$ shallower. Thus, given that S16 used the O3N2 calibrator applied to their CALIFA observations, this may explain the discrepancy between S16 and B17. We defer a more detailed comparison via synthetic observations of our simulations, tailored to each

observation, to future work. Rather, Fig. 2.3 provides a broad comparison, highlighting that the radial gradients of gas-phase [O/H] within our simulations lie within the scatter across the MW, M31, and nearby MW-mass galaxies.

Also, while we do not show it, we compared the [N/H] gradients of our sample at $z = 0$ to observations of M31 and the MW. Our mean [N/H] gradient is $-0.039\,\mathrm{dex\,kpc^{-1}}$, with a standard deviation of 0.007, and the full distribution across hosts spans -0.028 to $-0.057\,\mathrm{dex\,kpc^{-1}}$. This agrees well with values measured for the MW in Esteban & García-Rojas (2018) (comparing 3 different ionization correction factors, they found [N/H] gradients of -0.047 to $-0.050\,\mathrm{dex\,kpc^{-1}}$ with uncertainties of ≈ 0.008) and in Arellano-Córdova et al. (2020) (-0.049 ± 0.007). Our [N/H] gradients also agree with the value measured in M31 by Sanders et al. (2012) ($-0.0303 \pm 0.0049\,\mathrm{dex\,kpc^{-1}}$. However, [N/H] gradients measured in Rudolph et al. (2006) ($-0.071 \pm 0.010\,\mathrm{dex\,kpc^{-1}}$ in the optical and $-0.085 \pm 0.010\,\mathrm{dex\,kpc^{-1}}$ in the far infrared) and Fernández-Martín et al. (2017) ($-0.080 \pm 0.019\,\mathrm{dex\,kpc^{-1}}$) are substantially steeper than in our hosts. As with the [O/H] gradients, this may be because the gradients are not measured over the same radial range.

2.4.2. Evolution of radial gradients. We next explore the evolution of gas-phase radial gradients of [O/H], [Fe/H], and [N/H] at $z \leq 1.5$, over the last $\sim 10\,\mathrm{Gyr}$, during the primary epoch of disk assembly, to understand the initial conditions for star formation and chemical tagging of stars. In summary, we find that at earlier times the gas disk was more radially homogeneous (flatter gradients), so chemical tagging offers less discriminating power for radial birth location at earlier times.

Similar to Fig. 2.2, Fig. 2.4 shows radial profiles of [O/H], [Fe/H], and [O/Fe] in gas at different redshifts. The solid line shows the mean across our 11 galaxies, while the shaded region shows the 1-σ scatter. At all radii, [O/H] and [Fe/H] increase with time, as the gas mass declines while more stars enrich the Interstellar Medium (ISM). This evolution agrees with the observed gas-phase galaxy mass-metallicity relation (Tremonti et al., 2004).

Ma et al. (2016) explored this evolution across a wide galaxy mass range in the FIRE-1 simulations: they found that as galaxies grow more massive, the mass-loading factor of their winds decreases, and metals are more easily held in/near the galaxy as opposed to being driven into the halo (see also Muratov et al., 2015, 2017; Anglés-Alcázar et al., 2017).

Fig. 2.4 also shows that both [O/H] and [Fe/H] have negative radial gradients at all times. Ma et al. (2017a) also found primarily negative gradients in the FIRE-1 suite, because the high star-formation efficiency in the inner disks of galaxies with well ordered rotation leads to sustained negative radial gradients. At $z = 0$, our average change in [O/H] from $0 - 15\,\mathrm{kpc}$ is $\approx 0.51\,\mathrm{dex}$, while this is $\approx 0.56\,\mathrm{dex}$ for [Fe/H]. At $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$), the average change in abundance from $0 - 15\,\mathrm{kpc}$ is $\approx 0.24\,\mathrm{dex}$ for [O/H] and $0.28\,\mathrm{dex}$ for [Fe/H]. Furthermore, as expected given the scatter in formation history, we find larger host-to-host scatter (averaged over all radii) at earlier times: $0.09\,\mathrm{dex}$ for [O/H] and $0.07\,\mathrm{dex}$ for [Fe/H] at $z = 0$, but at $z = 1.5$ this was $0.2\,\mathrm{dex}$ for both elements.

Fig. 2.4 also shows that the abundance profiles were flatter (more homogeneous) at earlier times, because the abundance at smaller radii evolves more rapidly than at larger radii. Increased accretion/merger rates, coupled with higher star-formation rates and stronger gas turbulence, drove more efficient radial mixing at earlier times (Ma et al., 2017a). In FIRE simulations, early galaxies experience bursty, stellar feedback-driven outflows that radially mix the ISM, in addition to local turbulence. The profiles steepen with time, because as the gas disk settles down and becomes more rotationally supported, it is capable of sustaining stronger radial gradients given less radial mixing (Ma et al., 2017a). At $z \lesssim 1$ ($t_{\mathrm{lookback}} \lesssim 7.8\,\mathrm{Gyr}$), the radial profiles generally show an up-turn in the innermost bulge-dominated region. Beyond $\approx 4\,\mathrm{kpc}$, the profiles are well fit by a linear relation at all redshifts.

Fig. 2.4 (bottom panel) shows that [O/Fe] tends to decline over time at all radii, because at early times, core-collapse supernovae dominate the enrichment, which preferentially produce α-elements like O. However, the mean trends in Fig. 2.4 (bottom panel) are not

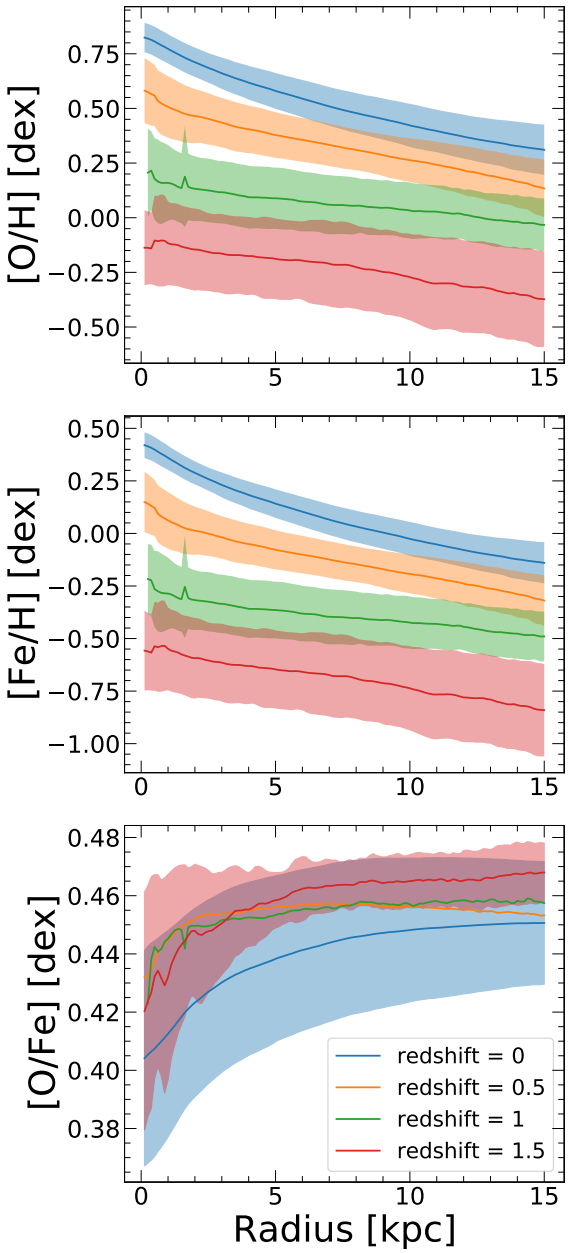


FIGURE 2.4. Radial profiles of gas-phase elemental abundances in our simulations at different redshifts. The solid lines show the mean and the shaded regions show the 1-σ scatter across our 11 galaxies. The top and middle panels show steepening radial profiles of [O/H] and [Fe/H] with time, as the gas disk becomes more rotation dominated, with less radial turbulence, thus sustaining stronger radial gradients. The bottom panel shows that the innermost regions of the gas disk have lower [O/Fe] than the outer disk, this indicates that the inner disk is more evolved, that is, has had more Ia supernovae that produce more Fe than α-elements (like O), than the outer disk.

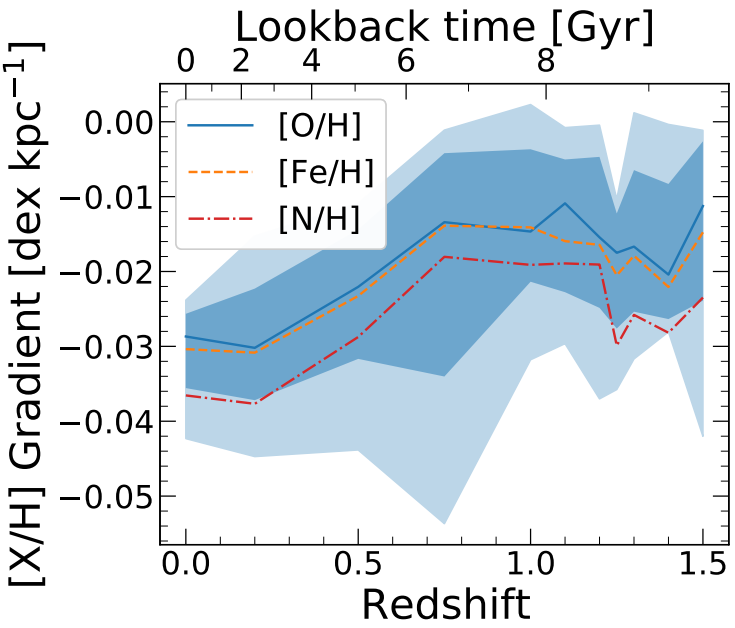


FIGURE 2.5. Evolution of radial gradients in gas elemental abundances. The lines show the means for [O/H], [Fe/H], and [N/H]. The dark shaded region shows the 1-σ scatter and the light shaded region shows the total scatter in [O/H] across our 11 galaxies. At each redshift we calculate the gradient via a linear fit across a redshift-dependent radial range: $4 < R < 12\,\mathrm{kpc}$ for $z < 1$ ($t_{\mathrm{lookback}} < 7.8\,\mathrm{Gyr}$) and $0 < R < 8\,\mathrm{kpc}$ for $z \geq 1$. The gradient is flattest ($\approx -0.015\,\mathrm{dex\,kpc^{-1}}$) at $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$), likely because high merger rates drive large scale radial turbulence, erasing the radial gradients of the disks. This leads to the azimuthally averaged metallicity at a given radius showing little radial dependence although stochastic enrichment/accretion events may lead to locally over or under-enriched regions. Shallow gradients persist until $z \approx 1$. At lower redshifts, the gas disk becomes more rotationally supported and is capable of sustaining stronger radial gradients, becoming steepest ($\approx -0.03\,\mathrm{dex\,kpc^{-1}}$) at $z = 0$. The gradients of [O/H] and [Fe/H] are almost identical, despite being sourced primarily by core-collapse and Ia supernovae, respectively. However, the [N/H] gradient, sourced primarily through stellar winds from massive stars, which have a metallicity dependent mass loss rate, is steeper at all times.

necessarily true for individual hosts. The dynamic range of [O/Fe] is small, so minor upticks in star formation rates lead to enhanced [O/Fe] as the rate of core collapse supernovae temporarily increases. Individual galaxies show overall increases in [O/Fe] following periods of increased star formation, and different radii of individual hosts show relative increases or decreases in [O/Fe], likely correlated with radial variations in star-formation rates.

On average, at later times, the (more delayed) Ia supernovae preferentially enrich the galaxy in Fe, driving down [O/Fe]. However, the typical change in gas-phase [O/Fe] at fixed radius from $z = 1.5$ ($t_{\rm lookback} = 9.4\,{\rm Gyr}$) to 0 is only $\approx 0.02 - 0.03\,{\rm dex}$. The [O/Fe] radial gradients are positive at all times, because the outer disk is always younger than the inner disk/bulge, though the gradients are weak at larger radii. The [O/Fe] radial profiles steepen at small radii at later times, at least at $z \lesssim 1$ ($t_{\rm lookback} \lesssim 7.8\,{\rm Gyr}$). Unlike the profiles of individual elements, the host-to-host standard deviation of [O/Fe] increases at later times, from $0.015\,{\rm dex}$ at $z = 1.5$ to $0.037\,{\rm dex}$ at $z = 0$. Overall, [O/Fe] does not provide strong discrimination power for chemical tagging at any radii or time that we examine.

While not shown here, we also measure the evolution of [Mg/Fe], which is more significant than [O/Fe]. In the outer disk ($R = 12\,{\rm kpc}$) [Mg/Fe] decreases from $\approx 0.3\,{\rm dex}$ at $z = 1.5$ to $\approx 0.22\,{\rm dex}$ at $z = 0$. In the inner disk ($R = 4\,{\rm kpc}$) the evolution is larger, from $\approx 0.29\,{\rm dex}$ to $0.18\,{\rm dex}$ over the same redshifts. The stronger evolution seen in [Mg/Fe] likely results from stellar winds in our simulations producing relatively little Mg. The stellar-wind model in FIRE-2 has a metallicity-dependent yield for O. At lower redshifts, as the gas disk becomes more enriched, O production also increases, leading to less evolution in [O/Fe].

Fig. 2.5 shows the evolution of the mean radial gradients of [O/H], [Fe/H], and [N/H] at redshifts 0, 0.2, 0.5, 0.75, 1, 1.1, 1.2, 1.25, 1.3, 1.4, and 1.5. The shaded regions show the 1-σ scatter. We also checked that the evolution of each galaxy qualitatively agrees with this mean evolution. We measure radial gradients between $4 < R < 12\,{\rm kpc}$ at $z < 1$ ($t_{\rm lookback} < 7.8\,{\rm Gyr}$), consistent with Fig. 2.3, and between $0 < R < 8\,{\rm kpc}$ at $z \geq 1$ ($t_{\rm lookback} \geq 9.4\,{\rm Gyr}$). (For a few hosts with bulge-like upturns in their profiles at $\lesssim 2\,{\rm kpc}$, we measure their gradients between $2 - 8\,{\rm kpc}$, where the gradient is nearly linear). We use a redshift-dependent radial range, because we are exploring the gas from which stars are forming. We select gas approximately contained in $R_{90}^{\rm star}$, which is $\approx 11\,{\rm kpc}$ at $z = 0$ and $\lesssim 8\,{\rm kpc}$ at $z \gtrsim 0.5$. At $z < 1$, we measure the gradient at $4 - 12\,{\rm kpc}$, because the inner

disk is dominated by the bulge region and has a steeper profile, as Fig. 2.4 shows. Thus, at late times we measure at $4 - 12\,\mathrm{kpc}$, which encompasses both the inner and outer disk and exhibits a nearly linear profile. We tested our analysis measuring the gradient over varying radial ranges and found that, while the normalization varies somewhat, the shape of the profile and the evolution are consistent, as Fig. 2.4.

As Fig. 2.5 shows, the strength of the radial gradient generally decreases over time. The minimum magnitude of the gradient is $\approx 0.01\,\mathrm{dex\,kpc^{-1}}$ and occurs at $z \approx 1.5$ ($t_{\rm lookback} \approx 9.4\,\mathrm{Gyr}$). We find just 2 galaxies that achieve a flat gradient at this time. At $z \lesssim 1$ ($t_{\rm lookback} \lesssim 7.8\,\mathrm{Gyr}$), the radial gradients gradually steepen to $-0.03\,\mathrm{dex\,kpc^{-1}}$ at $z = 0$. The gradients prior to $z \approx 1$ are approximately constant with redshift. While there are fluctuations in the gradient at high z, these are transient features in the simulations driven primarily through minor mergers, flybys of satellite galaxies, and starbursts. The host-to-host scatter is smallest at $z = 0$ and is largest at $z = 0.75$ ($t_{\rm lookback} = 6.6\,\mathrm{Gyr}$), in part because of one galaxy (m12f) that experiences a major merger at this time.

Fig. 2.5 also compares the evolution of the radial gradients in [O/H] and [Fe/H] with [N/H]. Consistent with most results in this paper, we find little-to-no difference between [O/H] and [Fe/H], despite their differing origins, from primarily core-collapse and primarily Ia supernovae, respectively. However, we find systematically stronger radial gradients in [N/H] at all times. Unlike O and Fe, which are sourced primarily through supernovae, N is sourced primarily by stellar winds in the FIRE-2 simulations, and the wind mass-loss rate from massive stars (in the first 3.5 Myr) depends roughly linearly on progenitor metallicity. (The N yield from core-collapse supernovae also increases linearly with progenitor metallicity in the FIRE-2 model, but this effect is subdominant, because most N comes from stellar winds.) The progenitor metallicity dependence of N (often called secondary production of N) results in enhanced N production in regions that are already more metal rich, and thus it drives a steeper gradient for N (by $\approx 0.015\,\mathrm{dex\,kpc^{-1}}$) than O or Fe at all times.

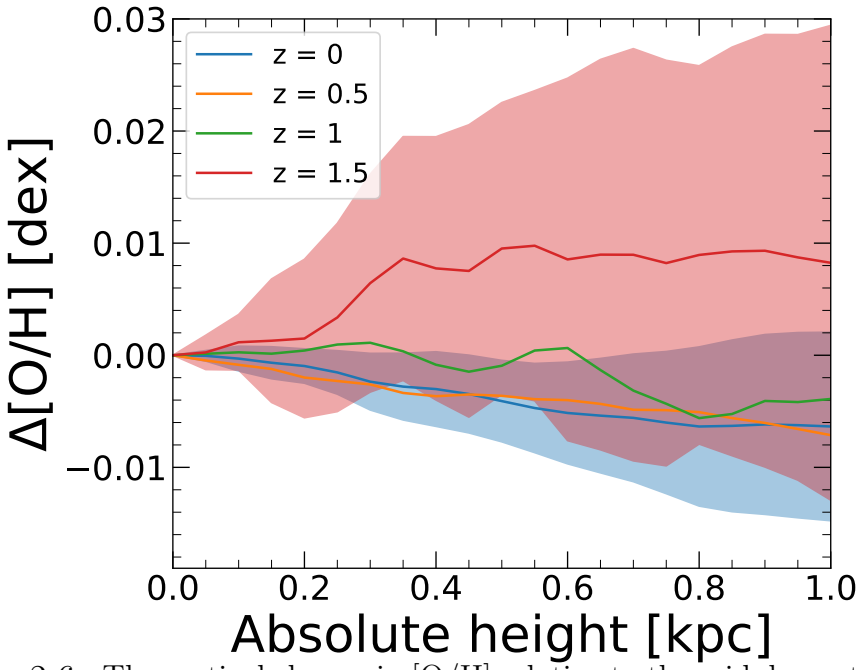


FIGURE 2.6. The vertical change in [O/H] relative to the midplane at various redshifts. We measure the vertical profile at $R = 7 - 9\,\mathrm{kpc}$. The solid lines show the mean while the shaded region shows the 1-σ standard deviation across 11 galaxies. [O/H] shows minimal variation with height: the mean deviation at 1 kpc above the midplane is $\approx 0.008\,\mathrm{dex}$ at $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$) and $\approx -0.006\,\mathrm{dex}$ at $z = 0$. The majority of stars form within a few hundred parsecs of the galactic midplane at low z and within $\sim 1.5\,\mathrm{kpc}$ of the midplane at the highest redshift we examine. Close to the plane of the disk ($|Z| \leq 200\,\mathrm{pc}$) the variation in [O/H] is even smaller at all times ($|\Delta[\mathrm{O/H}]| \lesssim 0.001\,\mathrm{dex}$ at $z = 1.5$ and $|\Delta[\mathrm{O/H}]| \lesssim 0.002\,\mathrm{dex}$ at $z = 0$). At $z < 1$ ($t_{\mathrm{lookback}} < 7.8\,\mathrm{Gyr}$) we find a weak but systematically negative gradient. We conclude that vertical gradients are not measurably significant for chemical tagging at any redshift.

2.4.3. Vertical profiles across time. We next examine the vertical profiles (in absolute height) of elemental abundances, for all gas near the solar circle, within a cylindrical radius of $7 < R < 9\,\mathrm{kpc}$. We normalize the vertical profiles by subtracting the midplane abundance at each redshift. Fig. 2.6 shows the vertical profiles for [O/H]. The solid line shows the mean and the shaded regions show the 1-σ scatter; we only show the scatter at $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$) and $z = 0$ for clarity.

Fig. 2.6 shows that any systematic trends in abundance with height to 1 kpc is $\lesssim 0.01\,\mathrm{dex\,kpc^{-1}}$ absolute on average at all times, and the 1-σ scatter is $\lesssim 0.01\,\mathrm{dex\,kpc^{-1}}$

at $z = 0$ and $\lesssim 0.02\,\mathrm{dex\,kpc^{-1}}$ at $z = 1.5$. Thus, the gas disk is well mixed vertically. In most of our galaxies, the deviations in abundance increase with distance from the midplane, that is, they shows a systematic gradient with height. Over time, these vertical gradients become increasingly (if weakly) more negative, which supports the idea of 'upside-down' disk formation (e.g. Bird et al., 2013; Ma et al., 2017a; Bird et al., 2021), such that stars formed in a more vertically extended disk at higher redshifts, leading to more enrichment at larger heights, at later times stars formed in a thinner disk and gas farther out of the midplane became relatively less enriched. At $z \gtrsim 1$, the absolute strength of this vertical gradient is in fact comparable to the radial gradient (Fig. 2.5), while at $z = 0$, the vertical gradient is $\approx 3\times$ weaker than the radial gradient. This is because the timescale for vertical mixing is short, given gas turbulence, and that the vertical scale-height of the gas is itself set by the maximum Jeans length at that time. Furthermore, with implications for chemical tagging, the majority of star formation in our simulations is limited to $\lesssim 500\,\mathrm{pc}$ of the midplane at $z < 0.5$, and $\lesssim 1.5\,\mathrm{kpc}$ up to $z < 1.5$, and Fig. 2.6 shows that vertical variations in abundance are minimal on those scales. The vertical trends in [Fe/H] (not shown here) are consistent with [O/H] within $\approx 0.01\,\mathrm{dex}$.

2.4.4. Azimuthal variations across time. We next investigate azimuthal variations of elemental abundances in gas, including its evolution. We thus test a common assumption in galactic evolution, that gas is well mixed azimuthally within a given annulus (e.g. Frankel et al., 2018, 2020).

Fig. 2.7 shows the standard deviation in [O/H] and [O/Fe] along angular bins of varying length at fixed radius. Specifically, we compute the standard deviation within a given angular bin, and Fig. 2.7 shows the mean standard deviation across all bins of a given angular size for all 11 simulations. We stack 3 snapshots ($\Delta t \approx 50\,\mathrm{Myr}$) for each redshift. We use an annulus of gas $\pm 0.3\,\mathrm{kpc}$ out of the plane of the disk because as shown in Fig. 2.6 gas-phase abundances are effectively homogeneous within this height. We also measure within a radius $\pm 0.15\,\mathrm{kpc}$ of

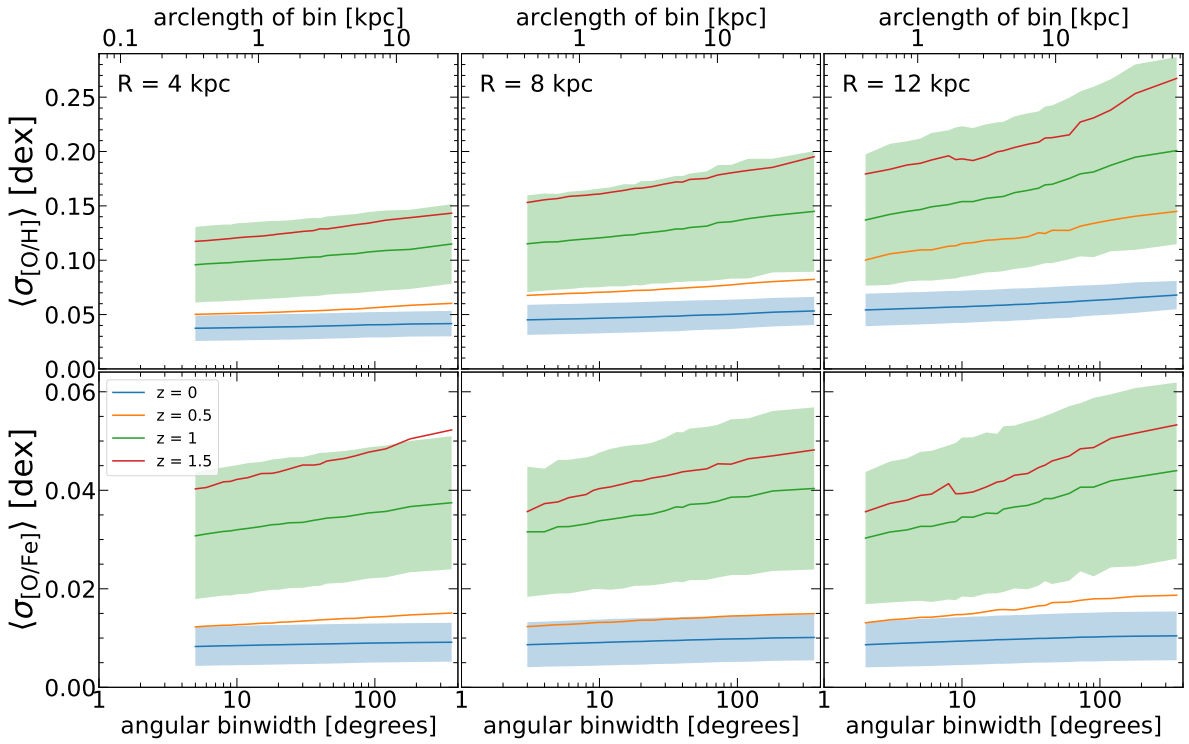


FIGURE 2.7. Azimuthal scatter in elemental abundances for [O/H] and [O/Fe] in gas, as a function of angular scale, at different redshifts and different radii. The solid lines show the mean and the shaded regions show the 1-σ scatter across our suite of 11 galaxies: we show scatter only at $z = 1$ ($t_{\rm lookback} = 7.8\,{\rm Gyr}$) and 0. The scatter increases as a function of angular bin size at all redshifts and at all radii. At $z = 0$, near the solar circle ($R = 8\,{\rm kpc}$), the average azimuthal scatter across the disk is $\approx 0.053\,{\rm dex}$ for [O/H] and [Fe/H] (not shown) and $\approx 0.009\,{\rm dex}$ for [O/Fe]. For all angular bin sizes, the average scatter increases with redshift: at earlier cosmic times the gas disks were less well mixed within a given annulus. At $z = 0$ the scatter across the disk is $\approx 0.2\,{\rm dex}$ for [O/H] and [Fe/H] (not shown) and $\approx 0.05\,{\rm dex}$ for [O/Fe]. The scatter also increases with angular bin size at all redshifts, although the increase is minimal at late times. At low z, this means that azimuthal variations are dominated by local (and not global) fluctuations in the disk. Finally, the azimuthal scatter increases with radius for individual abundances: gas is azimuthally better mixed in the inner disk, likely a result of shorter orbital times leading to faster mixing.

the selected cylindrical radius, to ensure that the angular length dominates over the radial length for our smallest angular bins, that is, to ensure that the radial gradient does not induce significant scatter. We show the 1-σ scatter for $z = 0$ and $z = 1$ ($t_{\rm lookback} = 7.8\,{\rm Gyr}$).

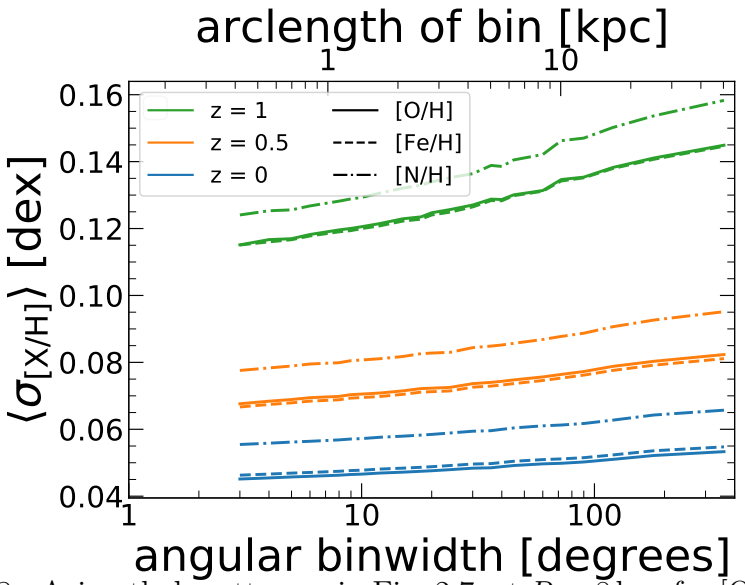


FIGURE 2.8. Azimuthal scatter, as in Fig. 2.7, at $R = 8\,\mathrm{kpc}$ for [O/H] (solid), [Fe/H] (dashed), and [N/H] (dash dot). As with the radial gradients, the azimuthal variations of [O/H] and [Fe/H] are almost identical, despite being sourced primarily by core-collapse and Ia supernovae, respectively, while the variation in [N/H], which is sourced primarily through metallicity-dependent stellar winds from massive stars, is higher at all redshifts.

We exclude m12c at $z = 1.5$ for angular scales $\Delta\phi \leq 8°$, because its angular bins contain too few gas particles.

At $z = 0$ and $R = 8\,\mathrm{kpc}$ (near the solar circle) the typical azimuthal scatter across the gas disk is $\lesssim 0.053\,\mathrm{dex}$ for [O/H], $\lesssim 0.055\,\mathrm{dex}$ for [Fe/H] (not shown), and $\lesssim 0.01\,\mathrm{dex}$ for [O/Fe]. This value for [O/H] agrees well with MW observations (Wenger et al., 2019) and observations of external galaxies (Sakhibov et al., 2018; Kreckel et al., 2019, 2020), though we emphasize that we are not measuring azimuthal scatter in the same way: those observations typically measure differences in abundances between arm and inter-arm regions or measure abundance variations between HII regions within an aperture of a given size.

Our azimuthal scatter decreases with smaller angular bin size, with a minimum of $\approx 0.045\,\mathrm{dex}$ for [O/H] ($\approx 0.046\,\mathrm{dex}$ for [Fe/H]) and $\approx 0.009\,\mathrm{dex}$ for [O/Fe]at the smallest angular scales. Interestingly, this minimal scatter remains well above 0 dex as $\Delta\phi$ goes to 0. We emphasize that our analysis does not zoom-in on Giant Molecular Clouds (GMC)

or individual star-forming regions, but rather we examine all of the ISM centered on (effectively) random positions. Thus, our results on small scales do not immediately inform the homogeneity of individual GMCs, especially given their short lifetimes ($\lesssim 7\,\mathrm{Myr}$) in our simulations (Benincasa et al., 2020), and we will examine GMC homogeneity in future work. Appendix. A.2 also examines how small-scale variations depend on our choice of diffusion coefficient for sub-grid turbulent mixing in gas.

The 1-σ host-to-host scatter is approximately independent of bin size and is $\lesssim 0.014\,\mathrm{dex}$ for [O/H], $\lesssim 0.015\,\mathrm{dex}$ for [Fe/H], and $\lesssim 0.005\,\mathrm{dex}$ for [O/Fe]. Thus, at $z = 0$ gas within all of our galaxies is well mixed, that is, the azimuthal scatter is comparable to typical measurement uncertainties ($\sim 0.05\,\mathrm{dex}$) for elemental abundances.

Fig. 2.7 shows that, at all radii and at all angular bin sizes, the azimuthal scatter was more significant at earlier times, that is, gas was less azimuthally mixed than it is now. This is likely because higher accretion and star formation rates combined with burstier star formation leads to more pronounced local pockets of enrichment in gas. At $R = 8\,\mathrm{kpc}$ and at $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$) the azimuthal scatter across the disk is $\lesssim 0.2\,\mathrm{dex}$ for [O/H]and [Fe/H] (not shown) and $\lesssim 0.05\,\mathrm{dex}$ for [O/Fe]. The scatter does not drop below $\approx 0.15\,\mathrm{dex}$ for either [O/H] or [Fe/H] at the smallest azimuthal scales ($0.035\,\mathrm{dex}$ for [O/Fe]). The 1-σ host-to-host scatter is $\lesssim 0.05\,\mathrm{dex}$ for [O/H] and [Fe/H] and $\lesssim 0.016\,\mathrm{dex}$ for [O/Fe].

Additionally, Fig. 2.7 shows that the difference in scatter between large and small angular scales varies with time. This difference is more significant at earlier times: at $z \gtrsim 1$ ($t_{\mathrm{lookback}} \gtrsim 7.8\,\mathrm{Gyr}$) this change is $\approx 0.042\,\mathrm{dex}$ at $R = 8\,\mathrm{kpc}$. Thus, at early times, galaxy-scale fluctuations are more important in driving azimuthal scatter (as visible in Fig. 2.1). However, at $z \sim 0$, the azimuthal scatter across small versus large angular scales differs by only $\approx 0.008\,\mathrm{dex}$, so small-scale fluctuations drive most of the azimuthal scatter (also visible

in Fig. 2.1). These results at low redshift are useful from an observational perspective, because they means that one can generalize smaller-scale observations of gas-phase abundances to overall azimuthal trends at fixed radius.

Fig. 2.7 also shows that the azimuthal scatter depends on radius. The azimuthal scatter increases with increasing radius for a given angular bin size, and in fact, this is true for both fixed angular and physical bin size. At $z = 0$ the azimuthal scatter across the entire disk at $R = 4\,\mathrm{kpc}$ is $\lesssim 0.042\,\mathrm{dex}$ for [O/H] ($\lesssim 0.046\,\mathrm{dex}$ for [Fe/H], not shown), and this increases to $\lesssim 0.062\,\mathrm{dex}$ for [O/H] and [Fe/H] at $R = 12\,\mathrm{kpc}$. At $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$) the azimuthal scatter ranges from $\lesssim 0.015\,\mathrm{dex}$ to $0.25\,\mathrm{dex}$ for [O/H] and [Fe/H]. We do not find radial dependence in [O/Fe], which has a maximal scatter $\lesssim 0.01\,\mathrm{dex}$ ($\lesssim 0.052\,\mathrm{dex}$) at all radii at $z = 0$ ($z = 1.5$). Most likely, the radial dependence in the azimuthal scatter of [O/H] and [Fe/H] results from the increase of the orbital timescale, and hence the timescale for mixing, with radius. Furthermore, cosmic accretion of under-enriched gas also likely contributes to this increase in azimuthal scatter with radius, especially at earlier times, when the increase with radius is stronger.

Fig. 2.8 shows the azimuthal scatter of [O/H], [Fe/H], and [N/H] at $R = 8\,\mathrm{kpc}$ at 3 redshifts. As we discussed above, metallicity-dependent stellar winds from massive stars, rather than supernovae, primarily source N, so this compares the azimuthal mixing of elements sourced by these different processes. The azimuthal scatter of [N/H] is larger than that of [O/H] and [Fe/H] for all bins at each redshift. On the scale of the entire annulus, the scatter in [N/H] is approximately $0.015\,\mathrm{dex}$ larger at $z = 1$ ($t_{\mathrm{lookback}} = 7.8\,\mathrm{Gyr}$) and approximately $0.01\,\mathrm{dex}$ larger at $z = 0$. This discrepancy is slightly smaller for smaller angular scales at $z = 1$ (approximately $0.01\,\mathrm{dex}$ difference), but the difference in azimuthal scatter is independent of scale at $z = 0$. As previously stated, our stellar-wind rate depends linearly on progenitor metallicity, which likely drives these differences for N as compared with O and Fe. One might expect Fe, being sourced primarily by Ia supernovae, to have less scatter at

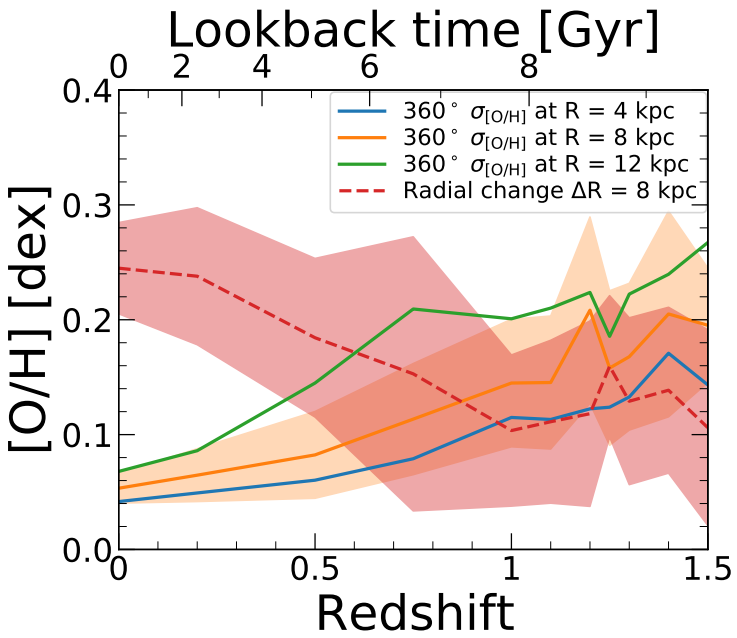


FIGURE 2.9. The evolution of variations in [O/H] in gas, both radially and azimuthally. (We do not show [Fe/H], its trends are consistent with [O/H] to $\lesssim 0.01$ dex.) The solid lines show the mean scatter, across our 11 galaxies, for the full (360°) annulus of gas at $R = 4$ kpc (blue), $R = 8$ kpc (yellow), and $R = 12$ kpc (green). The red dashed line shows the mean radial change in [O/H] across a radial distance of 8 kpc. The shaded regions show the 1-σ scatter. While the radial gradient dominates the spatial variations at late cosmic times, azimuthal variations were more significant than the radial gradient at earlier cosmic times ($z \gtrsim 0.9$, or $t_{\rm lookback} \gtrsim 7.4$ Gyr, at $R = 8$ kpc). Larger radii transition to radially dominated abundance variations at later times (see also Fig. 2.10). Thus, elemental evolution models should not assume azimuthal homogeneity of abundances at early times; instead, azimuthal variations are the primary source of spatial information for chemical tagging of stars forming at early times.

small scales than O, because Ia are caused by preferentially older stars than core-collapse supernovae, which occur closer to stellar birth location. A possible cause of our similarity comes from our assumed Ia delay-time distribution (Mannucci et al., 2006), which has a significant component from prompt Ia, at ages $\lesssim 100$ Myr. A Ia delay-time distribution with a more significant contribution from older stellar ages (Maoz & Graur, 2017) may lead to less small-scale scatter (e.g. Gandhi et al., 2022).

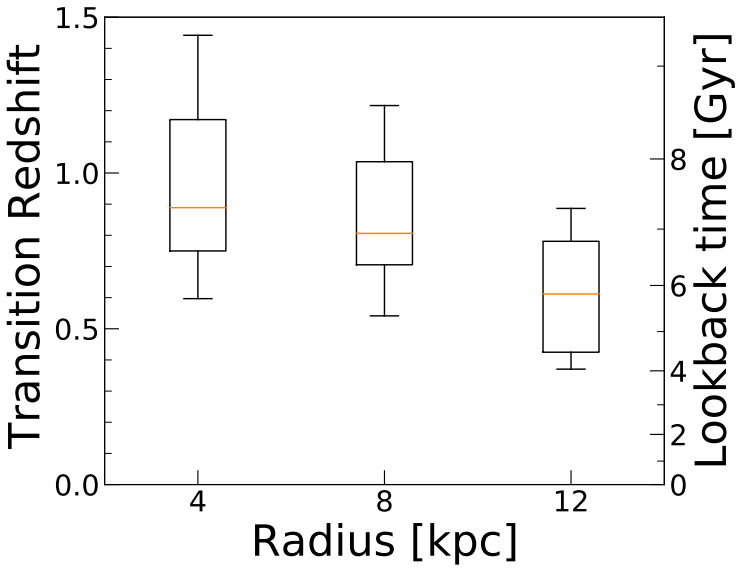


FIGURE 2.10. Following Fig. 2.9, the redshift below which radial variations in [O/H] dominate over azimuthal variations at 3 radii in our 11 simulations (the intersection of the dashed and solid lines). The horizontal line shows the median, the box shows the 68th percentile, and the whiskers show the full distribution. This transition redshift is the last time the azimuthal variation is stronger than the radial variation. This transition occurs earlier at smaller radii, where azimuthal variations are smaller (Fig. 2.7), given the shorter timescale for mixing at smaller radii. Before these transition redshifts, any model of chemical tagging should account for azimuthal variations as the primary source of spatial variation.

2.4.5. Azimuthal versus radial variations across time. We now compare the relative importance of radial gradients versus azimuthal scatter in gas-phase abundances. Fig. 2.9 shows the evolution of the radial variations in [O/H], from multiplying the radial gradient (as calculated in §2.4.2) of each simulation at each redshift by 8 kpc, to show the change from the inner to the outer disk. This measures the change from $4-12$ kpc at $z < 1$ ($t_{\rm lookback} < 7.8$ Gyr and from $0-8$ kpc at $z \geq 1$, where the radial profile is approximately linear (see Fig. 2.4). Fig. 2.9 also shows evolution of the azimuthal scatter around each disk (360°) at 3 radii.

Fig. 2.9 shows that at early times, $z \gtrsim 1$ ($t_{\rm lookback} \gtrsim 7.8$ Gyr), the 360° angular scatter in [O/H] at all radii is larger than the radial variation, with the angular scatter at large radii being approximately 2× higher. However, the radial variation dominates at all radii

at $z \lesssim 0.6$ ($t_{\rm lookback} \lesssim 5.8\,{\rm Gyr}$). At $z = 0$ the radial variations in [O/H] is approximately a factor of 4 higher than the azimuthal variation. Thus, at $z \sim 0$, one can approximate the gas disk variations primarily via the radial gradient, but at earlier times, the azimuthal variations dominate.

Fig. 2.10 shows the redshift (and lookback time) when the radial variation begins to dominate over the azimuthal scatter in [O/H] (the intersection of the solid and dashed lines in Fig. 2.9), for 3 radii. We measure this transition redshift separately for each simulation: the horizontal lines show the median, and the boxes and whiskers show the 68th percentile and the full distribution. The radial variation starts to dominate earlier at smaller radii, because the azimuthal scatter increases with radius at all redshifts (Fig. 2.7). The median transition redshift, after which the radial variation dominates, is $z \approx 0.9$ ($t_{\rm lookback} \approx 7.4\,{\rm Gyr}$) at $R = 4\,{\rm kpc}$, $z \approx 0.8$ ($t_{\rm lookback} \approx 6.9\,{\rm Gyr}$) at $R = 8\,{\rm kpc}$, and $z \approx 0.6$ ($t_{\rm lookback} \approx 5.8\,{\rm Gyr}$) at $R = 12\,{\rm kpc}$.

While we in general find no systematic differences between our LG-like hosts and our isolated hosts, the transition redshift of the LG-like simulations is systematically higher than that of the isolated simulations at large radii. At $R = 12\,{\rm kpc}$ the the median transition of the LG suite occurs at $z \approx 0.8$ ($t_{\rm lookback} \approx 6.9\,{\rm Gyr}$) compared to $z \approx 0.5$ ($t_{\rm lookback} \approx 5.1\,{\rm Gyr}$) for the isolated hosts. The overall distribution of transition redshifts for the isolated hosts falls within the distribution of transition redshifts for the LG suite, with the 68th percentiles almost entirely overlapping. The slight tendency for LG-like hosts to transition to radial domination earlier is unsurprising, given the results in Santistevan et al. (2020), who found that the main progenitors of our LG-like galaxies formed earlier than those of the isolated galaxies. This earlier formation may be responsible for the earlier transition redshift as the disks undergo more orbital times, which smooths azimuthal variation, and have a longer time to build up strong radial gradients.

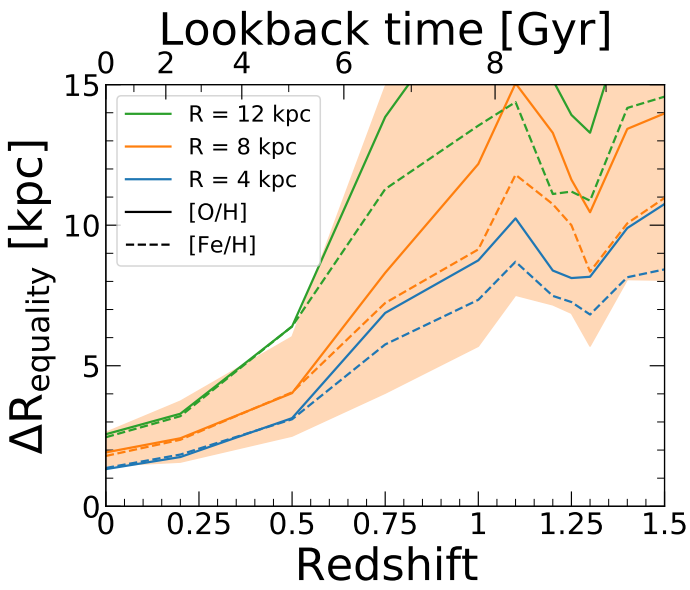


FIGURE 2.11. $\Delta R_{\rm equality}$ is the ratio of the azimuthal scatter to the radial gradient: it indicates the characteristic radial scale below which azimuthal variations dominate over radial variations, for [O/H] (solid) and [Fe/H] (dashed). We compute the azimuthal scatter using the full (360°) annulus of gas at $R = 4$, 8, and 12 kpc. Lines show the median, and shaded regions show the 68th percentile across our 11 galaxies. At early times, $\Delta R_{\rm equality}$ is comparable to the size of the gas disk, when azimuthal scatter dominated over radial variation, but after $z \approx 0.5$ ($t_{\rm lookback} = 5.1$ Gyr) the disk is well mixed azimuthally. At $z = 0$ the median $\Delta R_{\rm equality}$ near the solar circle is ≈ 1.8 kpc for [O/H] and ≈ 1.7 kpc for [Fe/H], while at $z = 1.5$ ($t_{\rm lookback} = 9.4$ Gyr) it is ≈ 14 kpc for [O/H] and ≈ 11 kpc for [Fe/H]. For radial scales less than $\Delta R_{\rm equality}$, the primary source of inhomogeneity of elemental abundances in gas is azimuthal variations, which complicate/limit the use of elemental abundances to infer a star's birth location to a radial precision smaller than $\Delta R_{\rm equality}$.

Fig. 2.11 shows another metric for comparing the importance of azimuthal versus radial variations across time. For each simulation, we compute $\Delta R_{\rm equality}$, the ratio of the 360° azimuthal scatter (at a given radius) to the radial gradient. This represents the characteristic radial scale over which the radial and azimuthal abundance variations are equal. The lines show the median $\Delta R_{\rm equality}$ for [O/H] and [Fe/H], measuring the azimuthal variation at 3 radii, and the shaded region shows the 68th percentile of [O/H] at $R = 8$ kpc. We apply

boxcar smoothing to the data to minimize the significant stochasticity in these values at early times, when the radial gradient fluctuated over short timescales (see Sec. 2.4.2).

The median $\Delta R_{equality}$ for [O/H] is $\lesssim 14\,kpc$ at $z = 1.5$ ($t_{lookback} = 9.4\,Gyr$) and $\lesssim 1.8\,kpc$ at $z = 0$ at the solar circle. This corresponds to the radial range over which azimuthal scatter dominates the variations in abundance, rather than the radial gradient. Thus, for the purposes of chemical tagging, this represents a limit for the radial precision that chemical tagging (of a single element) can place on the formation location of a star without also modeling azimuthal location. At early times, $\Delta R_{equality}$ is comparable to or greater than the size of the disk, meaning that the azimuthal coordinate determines the abundance of newly forming stars more than the radial position.

$\Delta R_{equality}$ is largest at $z = 1.5$ ($t_{lookback} = 9.4\,Gyr$) and then decreases with time, given the decreasing azimuthal scatter and increasing strength of the radial gradient with time. Also, over time the scatter across our 11 galaxies decreases. The high scatter at high redshifts is a result of the large scatter in both radial gradients and azimuthal variations at these times, given scatter in formation history.

Fig. 2.11 also shows the dependence of $\Delta R_{equality}$ on radius. This ratio slightly increases as a function of radius, because the azimuthal scatter increases with radius at all times (Fig. 2.7). This means that modeling chemical tagging of stellar birth radius is more challenging at larger radii.

In summary, any models for chemical tagging should incorporate azimuthal scatter in abundance especially at $z \gtrsim 0.6$ ($t_{lookback} \gtrsim 5.8$), because the azimuthal scatter in gas dominates at these early times.

2.4.6. Radial scale of measurable homogeneity. We next explore observational implications of our measured radial gradients, by comparing them against typical measurement uncertainties in elemental abundances, to understand the radial scales over which gas is effectively homogeneous in a measurable sense. Thus, in this sub-section we ignore azimuthal

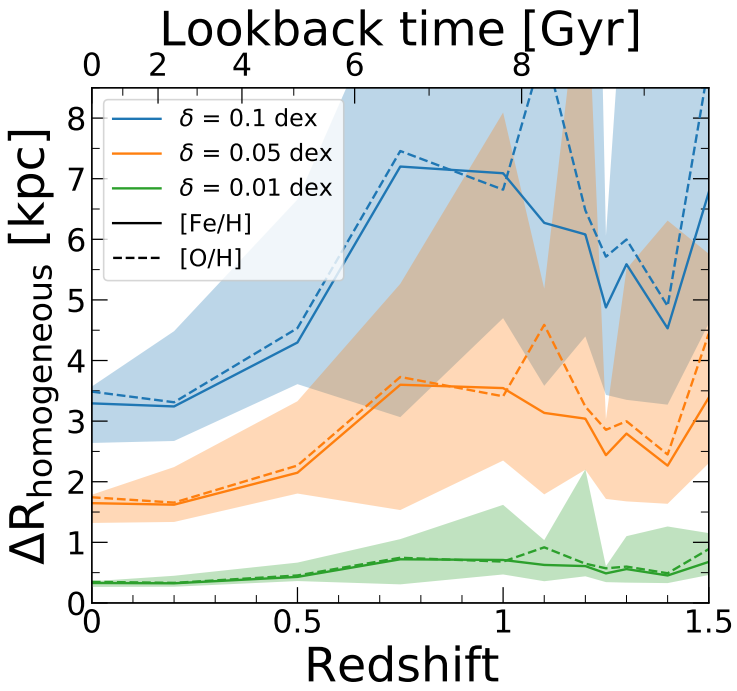


FIGURE 2.12. Assuming a given measurement uncertainty in an elemental abundance, $\Delta R_{\rm homogeneous}$ is the ratio of this measurement uncertainty to the radial gradient in that abundance; it is the characteristic radial scale below which the gas disk is effectively radially homogeneous in a measurable sense. We show $\Delta R_{\rm homogeneous}$ versus redshift, for 3 measurement uncertainties, motivated by observational surveys of stellar abundances. The lines show the median across our 11 galaxies for [Fe/H] (solid) and [O/H] (dashed), and the shaded regions show the 68th percentiles for [Fe/H]. At $z = 1.5$ ($t_{\rm lookback} = 9.4\,{\rm Gyr}$) the gas disk is measurably homogeneous (for $\delta = 0.05\,{\rm dex}$) across significant radial scales, $\Delta R \lesssim 3\,{\rm kpc}$, though this decreases to $\lesssim 1.6\,{\rm kpc}$ at $z = 0$. For high-precision abundance measurements ($\delta = 0.01\,{\rm dex}$), the gas disk is measurably homogeneous at $\Delta R \lesssim 0.6\,{\rm kpc}$ at $z = 1.5$ and $\Delta R \lesssim 0.3\,{\rm kpc}$ at $z = 0$. This highlights the limitations from just measurement uncertainties on chemical tagging to infer a star's birth radius; it does not include the additional complications from azimuthal variations (Fig. 2.11), which can be more important.

variations and focus just on radial gradients. While we examine gas-phase abundances, the chemical tagging of stars (that form out of this gas) ultimately motivates our work, so we examine measurement uncertainties typical of MW stellar surveys.

Motivated by observational surveys of stellar abundances, we select 3 observational measurement uncertainties of $\delta_{\rm m} = 0.01$, 0.05, and 0.1 dex. Fig. 2.12 shows $\Delta R_{\rm homogeneous}$, the

ratio of $\delta_{\rm m}$ to the radial gradient in abundance, versus redshift. Unlike Fig. 2.11, which shows radial scales of homogeneity at different radii, which depends on the azimuthal scatter at each radius, Fig. 2.12 depends only on the radial gradient measured across the whole disk. Thus the radial scale in Fig. 2.12 represents the average radial scale of observed homogeneity based on an azimuthally averaged radial gradient. The solid line shows the median and the shaded region shows the 68th percentile across our 11 hosts. $\Delta {\rm R}_{\rm homogeneous}$ represents the radial scale of measurable homogeneity: below this radial length scale, the change in abundance from the radial gradient is less than this measurement uncertainty. The larger $\Delta {\rm R}_{\rm homogeneous}$ is, the less precisely measurements can pinpoint a star's birth radius.

For a fiducial measurement uncertainty of $\delta = 0.05\,{\rm dex}$ at $z = 1.5$ ($t_{\rm lookback} = 9.4\,{\rm Gyr}$) $\Delta {\rm R}_{\rm homogeneous} \approx 3.1\,{\rm kpc}$, which drops to $\approx 1.6\,{\rm kpc}$ at $z = 0$ for [Fe/H]. [O/H] is consistent to within $\approx 1\,{\rm kpc}$ at all redshifts, except for $z = 1.5$. At early times, $\Delta {\rm R}_{\rm homogeneous}$ is larger and has large scatter. The largest scatter, at $z = 1.5$, comes from the radial gradients being flattest: some galaxies have gradients approaching $0\,{\rm dex\,kpc^{-1}}$, as Fig. 2.5 shows. Fig. 2.12 shows essentially an inverse gradient, so the short time fluctuations at early times (see Sec. 2.5) also lead to rapid and significant variations in $\Delta {\rm R}_{\rm homogeneous}$. After $z = 0.75$ ($t_{\rm lookback} = 6.6\,{\rm Gyr}$) $\Delta {\rm R}_{\rm homogeneous}$ decreases over time. This means that chemical tagging with measured abundances can identify the birth radius of more recently formed stars more precisely.

Comparing Fig. 2.11 with Fig. 2.12 shows that, in terms of limitations on chemical tagging for a star's birth radius, at $z \gtrsim 0.5$ ($t_{\rm lookback} \gtrsim 5.1\,{\rm Gyr}$) azimuthal variations dominate over observational uncertainties in the inner disk, for a fiducial uncertainty of $\delta = 0.05\,{\rm dex}$. In the outer disk ($R \geq 8\,{\rm kpc}$) azimuthal variations are larger than observational uncertainties for all redshifts. For higher-precision measurements, $\delta = 0.01\,{\rm dex}$, azimuthal variations dominate at all times at all radii. This implies that, if a primary motivation for chemical tagging is inferring the birth location of a star, there is not much benefit in pushing to higher precision,

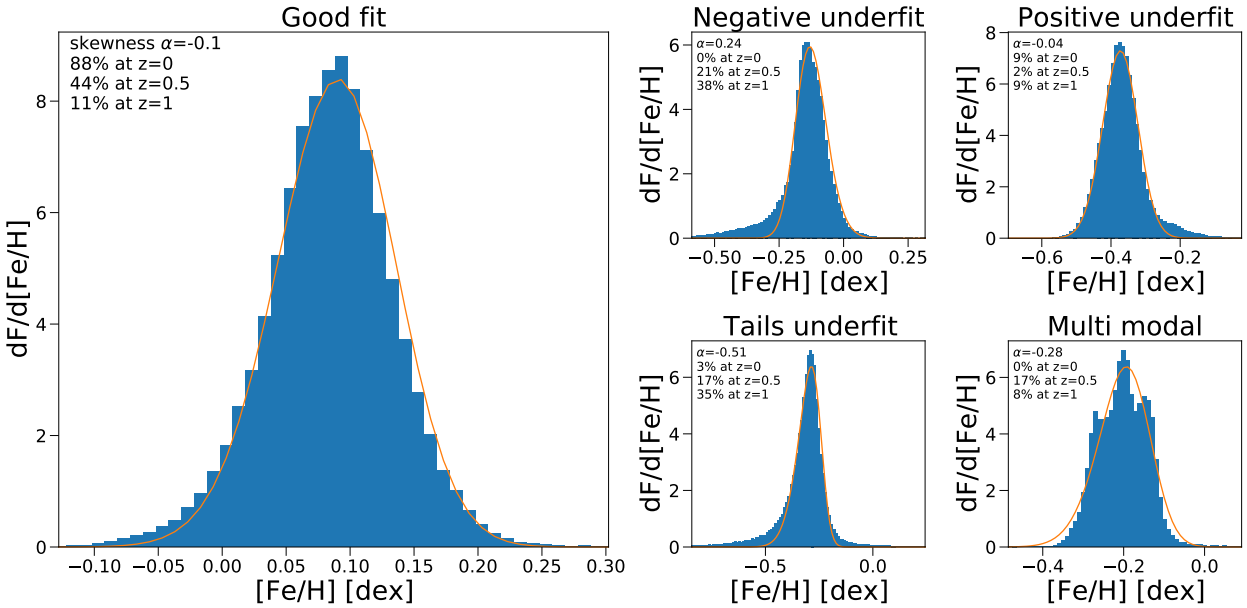


FIGURE 2.13. Example distributions of [Fe/H] from our 11 galaxies. Each panel shows the elemental distribution for a single simulation at a single radius. For each distribution, we stack 3 consecutive snapshots ($\Delta t \approx 50\,\mathrm{Myr}$) and measure all gas within $\pm 0.2\,\mathrm{kpc}$ of 3 radii ($R = 4$, 8, and 12 kpc) and within a height $\pm 1\,\mathrm{kpc}$ of the disk. The solid line shows a skew normal fit to the distribution. The left panel shows a well fit skew normal distribution. The right panel shows typical failure modes of fitting a skew normal distribution: underfitting negative skewness, underfitting positive skewness, underfitting the tails of the distribution, and multi modal distributions. Each panel shows the fitted skewness for the example distribution along with the percentage of fits that fall into the category at each redshift across all radii and for both [O/H] and [Fe/H]. In general, the simulations are well fit by a skew normal at $z = 0$, but it provides a worse fit at higher redshift, when the (negative) tails are preferentially underfit. In general, the fit to the distribution of [O/H] and [Fe/H] for the same R and redshift fall into the same category.

because azimuthal variations dominate. In fact, Fig. 2.9 can indicate the maximum precision in elemental abundance that one should aim to measure stars of a given age for this purpose, given our predicted azimuthal scatter, unless a given chemical tagging approach includes modeling azimuthal variations. We will explore possible models in future work.

2.4.7. Distributions of elemental abundances. Finally, we explore the full distributions of elemental abundances in gas that our simulations predict. Again, we emphasize

that our FIRE-2 simulations explicitly model the sub-grid diffusion/mixing of elemental abundances in gas via unresolved turbulent eddies, which is necessary to match observed distributions of abundances in the local Universe (Su et al., 2017; Escala et al., 2018; Hopkins et al., 2018).

We measure [O/H] and [Fe/H] distribution at $R = 4\,\mathrm{kpc}$, $R = 8\,\mathrm{kpc}$, and $R = 12\,\mathrm{kpc}$ at $z = 1$ ($t_{\rm lookback} = 7.8\,\mathrm{Gyr}$), $z = 0.5$ ($t_{\rm lookback} = 5.1\,\mathrm{Gyr}$), and $z = 0$ for all galaxies. We fit these with a skew normal distribution, using the LevMarLSQ fitter in Astropy (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018):

$$\frac{dF}{dx} = A \times \exp\left(-0.5\left(\frac{x-\mu}{\sigma}\right)^2\right) \times \frac{1 + erf\left(\alpha \times \frac{x-\mu}{\sqrt{2}\sigma}\right)}{2} \tag{2.1}$$

where μ is the mean, σ is the standard deviation, and α is the skewness. Fig. 2.13 shows representative example distributions of [Fe/H], for good and bad fits to this distribution, for a single simulation, and we list the percent of galaxies and radii that fall into each category.

As the left panel of Fig. 2.13 shows, a skew normal distribution reasonably fits these distributions in most cases at $z \sim 0$. However, there are several common failure modes. We categorize them as: failing to capture the positive or negative tails of the distribution, failing to capture the width of the distribution, or the distribution being multi-modal. The right panel of Fig. 2.13 shows examples of each of these failures, along with the percentage of fits ([Fe/H] and [O/H] combined) that we identify to fall into each category at each redshift, stacking all galaxies and radii at that redshift. In general, the fit to [O/H] and [Fe/H] at a given redshift and radius falls into the same category. At $z = 0$, the vast majority ($\approx 88\%$) of the distributions are well fit. The most common failure is a positive underfit, given pockets of high metal enhancement from feedback. However, at $z = 1$ ($t_{\rm lookback} = 7.8\,\mathrm{Gyr}$) the failures are more common, and only $\approx 11\%$ are well fit. Most common is having a negative underfit or both tails underfit, likely driven by more rapid accretion of low-metallicity gas at earlier times.

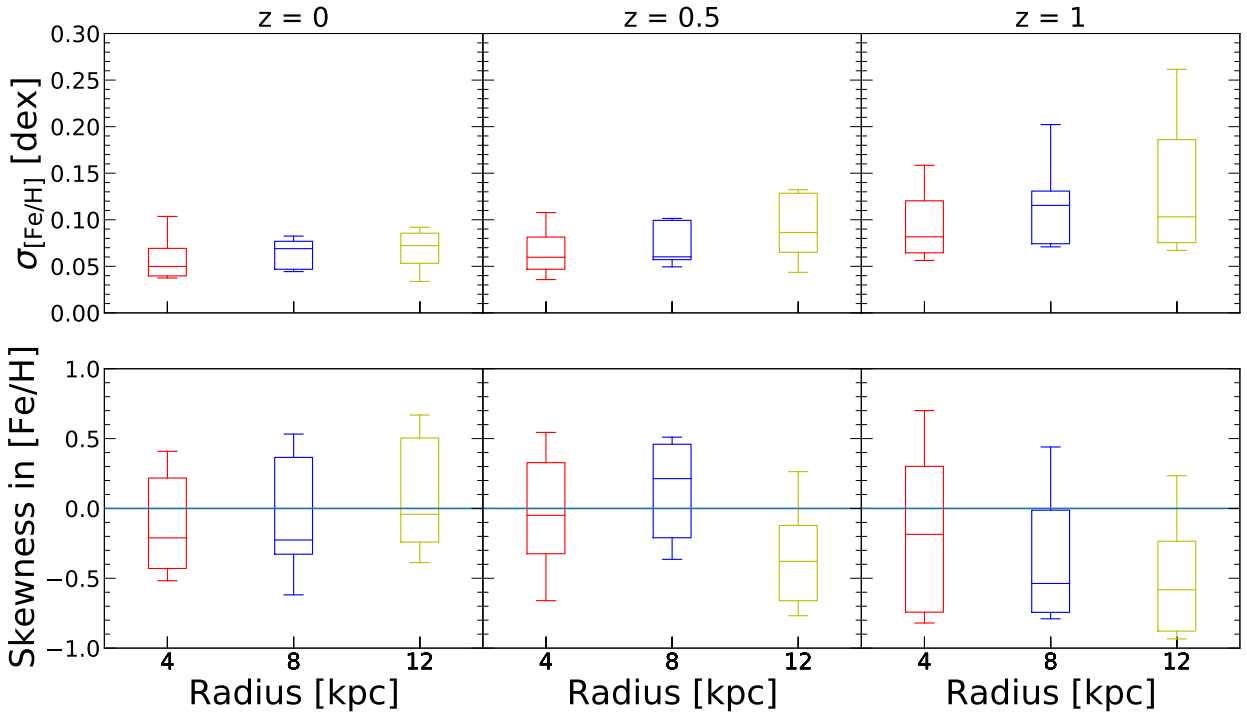


FIGURE 2.14. **Top**: standard deviation of the fitted elemental distribution of gas [Fe/H] at 3 radii, for increasing redshift (left to right). For each distribution, we stack 3 consecutive snapshots ($\Delta t \approx 50\,\mathrm{Myr}$) and measure all gas within $\pm 0.2\,\mathrm{kpc}$ of each radius and within a height $\pm 1\,\mathrm{kpc}$ of the disk. Each box shows the 68th percentile of the distribution and the whiskers show the full distribution of standard deviations. The standard deviation increases with radius and decreases with time. **Bottom**: skewness of the fitted elemental distributions of gas [Fe/H]. At earlier times, the gas disk had stronger negative skewness, but the disk relaxes to near zero skewness at $z = 0$. The skewness shows a slight radial dependence at both $z = 1$ ($t_{\mathrm{lookback}} = 7.8\,\mathrm{Gyr}$) and $z = 0$. At $z = 1$ the distributions at larger radii were more negatively skewed, whereas at $z = 0$ the distributions at smaller radii are more negatively skewed.

While not perfect, especially at earlier times, a skew normal fit does provide a simple characterization of the full distribution. Thus, Fig. 2.14 shows the fit parameters for [Fe/H] ([O/H], not shown, is consistent with this) at different radii and redshifts. The box-and-whisker plots show the median, 68th percentile, and the full distribution. The top row shows the fitted standard deviation, while the bottom row shows the fitted skewness, and left to right shows increasing redshift.

At all radii, σ decreases over time. At $R = 8\,\mathrm{kpc}$, near the solar circle, the median standard deviation decreases from $\approx 0.12\,\mathrm{dex}$ at $z = 1$ ($t_{\mathrm{lookback}} = 7.8\,\mathrm{Gyr}$) to ≈ 0.07 at $z = 0$ for [Fe/H] (≈ 0.11 to ≈ 0.06 for [O/H]). Thus, consistent with the results for azimuthal scatter, gas at a given radius becomes more homogeneous over time. Also consistent with our results for azimuthal scatter, Fig. 2.14 shows that at all redshifts, σ increases with radius, that is, metals are less well mixed at larger radii. At $z = 0$ the median is $\sigma \approx 0.05$ at $R = 4\,\mathrm{kpc}$ and increases to ≈ 0.07 at $R = 12\,\mathrm{kpc}$ for [Fe/H] (≈ 0.04 to ≈ 0.06 for [O/H]).

Fig. 2.14 (bottom row) shows that the distributions are preferentially negatively skewed at earlier times, but they trend toward Gaussian over time. At $R = 8\,\mathrm{kpc}$ the median skewness is $\alpha \approx -0.54$ at $z = 1$ ($t_{\mathrm{lookback}} = 7.8\,\mathrm{Gyr}$) and $\alpha \approx -0.21$ at $z = 0$ for [Fe/H] ($\alpha \approx -0.52$ to ≈ -0.23 for [O/H]). At earlier times, skewness decreases with radius, from ≈ -0.19 at $R = 4\,\mathrm{kpc}$ to ≈ -0.58 at $z = 1$ (≈ -0.15 to ≈ -0.5 for [O/H]). At earlier times, higher rates of cosmic accretion of pristine gas can skew the distributions negatively, especially at large radii, where enrichment also is more stochastic given lower star-formation rates and orbital/mixing times are longer. At later times, as the gas accretion and star-formation rates decrease, the distributions tend toward Gaussian, as metals become better mixed within each annulus. At $z \sim 0$, all radii show abundance distributions consistent with no skewness at the 1-σ level.

2.5. Summary and Discussion

2.5.1. Summary. We use a suite of FIRE-2 cosmological zoom-in simulations of 11 MW/M31-mass galaxies to explore the 3-D spatial variations and evolution of elemental abundances [O/H], [Fe/H], and [N/H] in gas at $z \leq 1.5$ ($t_{\mathrm{lookback}} \leq 9.4\,\mathrm{Gyr}$), to understand the birth conditions of stars to inform the efficacy of chemical tagging of stars. While many stars form prior to $z = 1.5$, the last $\sim 10\,\mathrm{Gyr}$ mark the primary epoch of disk assembly, which is where we are primarily interested in chemically tagging stars. Our main results are:

- *Vertical gradients*: are negligible. Abundances in gas are well mixed vertically at all times. At $R = 8\,\mathrm{kpc}$, the mean deviation in [O/H] at 1 kpc from the galactic midplane is $< 0.01\,\mathrm{dex}$ at all times. The inner $\sim 200\,\mathrm{pc}$ of the disks, where the majority of star formation for $z < 0.5$ occurs, is approximately uniform in abundance ($|\Delta[\mathrm{O/H}]| \lesssim 0.002\,\mathrm{dex}$) at all times. The inner $\sim 1.5\,\mathrm{kpc}$ of the disks, where the majority of star formation for $z > 0.5$ occurs has minimal vertical variation in abundance $|\Delta[\mathrm{O/H}]| \lesssim 0.01\,\mathrm{dex}$ at all times. Thus there is minimal vertical information for chemical tagging.
- *Radial gradients*: are negative at all times and for all abundances, with a maximum steepness of $\approx -0.03\,\mathrm{dex\,kpc^{-1}}$ at $z = 0$ and a minimum of $\approx -0.01\,\mathrm{dex\,kpc^{-1}}$ at $z \gtrsim 1$ ($t_{\mathrm{lookback}} \gtrsim 7.8\,\mathrm{Gyr}$). Radial gradients become steeper over time, because the disks become more rotationally supported and are better able to sustain a gradient against radial mixing, as noted in analysis of FIRE-1 simulations in Ma et al. (2017a). [N/H] has a steeper gradient at all times, because their production is dominated by stellar winds, whose mass-loss rates increase with metallicity in our simulations, enhancing the discrepancies between metal-rich and metal-poor regions. [O/Fe] shows little variation with redshift, and is approximately flat across the disk indicating it provides limited discriminating power for chemically tagging birth radii. Our [O/H] gradients broadly agree with most observations of nearby MW-mass galaxies, including M31, at the 1 or 2-σ level, though our gradients are somewhat steeper on average. By contrast, our gradients are somewhat shallower than most observations of the MW, though they agree at the 1 or 2-σ level with 9 of 13 MW observations.
- *Azimuthal scatter*: systematically decreases over time for all abundances, from $\approx 0.2\,\mathrm{dex}$ at $z = 1.5$ ($t_{\mathrm{lookback}} = 9.4\,\mathrm{Gyr}$) to $\approx 0.05\,\mathrm{dex}$ at $z = 0$ for [O/H] and [Fe/H] around the entire disk at $R = 8\,\mathrm{kpc}$. This evolution is a result of higher gas

accretion and also star-formation rates at earlier times, which lead to stronger variations in abundances on small scales, especially at larger radii, where orbital/mixing timescales are longer. The azimuthal scatter in [N/H] is larger (by ≈ 0.01 dex at $R = 8$ kpc) at all times than in [Fe/H] or [O/H], for the same reasons as above. Azimuthal variations reduce somewhat with smaller azimuthal aperture. However, even in angular bins as small as ≈ 350 pc, they remain ≈ 0.04 dex at $z = 0$ and ≈ 0.1 dex at $z = 1$ ($t_{\rm lookback} = 7.8$ Gyr) for [O/H] and [Fe/H]. We emphasize that our azimuthal bins do not center on GMCs or star-forming regions, so our results probe the homogeneity of effectively random patches of gas. We find good agreement between our azimuthal scatter in [O/H] in gas at $z = 0$ (≈ 0.05 dex) and observations of nearby galaxies (Sakhibov et al., 2018; Kreckel et al., 2019, 2020).

- *Azimuthal versus radial scatter*: At early times, the azimuthal scatter was larger than the radial variation for all abundances. We quantify the redshifts when the radial variation (across $\Delta R = 8$ kpc) first dominates over the azimuthal scatter, finding a median of $z \approx 0.9$ ($t_{\rm lookback} \approx 7.4$ Gyr) at $R = 4$ kpc and $z \approx 0.6$ ($t_{\rm lookback} \approx 5.8$ Gyr) at $R = 12$ kpc. Before this time, stars born at the same radius could have the same difference in metallicity as stars born $\Delta R \gtrsim 8$ kpc apart. We also quantify across time the radial range over which the radial and azimuthal variations are comparable, $\Delta R_{\rm equality}$. At $z \sim 0$, $\Delta R_{\rm equality}$ is small at ≈ 1.8 kpc, but at $z \gtrsim 1$ ($t_{\rm lookback} \gtrsim 7.8$ Gyr) $\Delta R_{\rm equality}$ is larger than the size of the disk. These results indicate that azimuthal variations in abundances provide the *dominant* information content for chemical tagging for stars formed $\gtrsim 6$ Gyr ago, so future approaches to chemical tagging of stars should start to incorporate/model these significant azimuthal variations.

- *Measurable homogeneity*: We quantified the radial scales across which our gas disks are effectively homogeneous in a measurable sense, given representative measurement uncertainties. For an uncertainty in elemental abundance of 0.05 dex, our gas disks are measurably homogeneous across $\Delta R \approx 1.7\,\mathrm{kpc}$ at $z = 0$ and $\Delta R \approx 3.5\,\mathrm{kpc}$ at $z \gtrsim 0.75$ ($t_{\mathrm{lookback}} \gtrsim 6.6\,\mathrm{Gyr}$). Moreover, azimuthal variations at $R \gtrsim 8\,\mathrm{kpc}$ are larger than 0.05 dex at all times. Thus, for any measurement uncertainty at or below this, using chemical tagging to measure birth radius is limited not by measurement uncertainty but instead by azimuthal variations. These results inform the needed precision for observations, given targeted precision for chemical tagging of stars across age/time. For example, if one only cares about modeling birth radius, there is little-to-no benefit in measuring a stellar abundance to better than ≈ 0.05 dex.
- *Elemental abundance distributions*: We measured the full distributions of elemental abundances in radial annuli and fit them with skew normal distributions. The skew normal distributions fit these distributions reasonably well, but there are failure modes that become more common at higher redshift, most notably underfitting the negative tails of the distribution and simultaneously underfitting the positive and negative tails. We find typically negatively skewed normal distributions at $z \gtrsim 1$ ($t_{\mathrm{lookback}} \gtrsim 7.8\,\mathrm{Gyr}$), with stronger negative skewness at larger radii. The distributions evolve toward approximately Gaussian distributions at all radii by $z = 0$.

2.5.2. Discussion. The primary goal of this paper is to understand the homogeneity of gas as a proxy for the birth conditions of stars across space and time, as a first step to understanding the efficacy of chemical tagging in a cosmological context. There are caveats to our analysis, though. Namely, our analysis is performed looking at individual elements, with the exception of [O/Fe]. Examining multi-element abundance distributions may well offer more discerning power. Our analysis of [O/Fe] suggests that this may be

limited. Furthermore, uncertainty in our fiducial diffusion coefficient leads to uncertainty in our small scale azimuthal abundance scatter, as seen in Appendix. A.3.

Additional complications arise when comparing our simulations to observational data. We present results in the context of constraining chemical tagging in the MW, but our simulations are not exact MW-analogs. Also, when comparing the redshift evolution of our results to observations of external galaxies, we track the evolutionary history of individual galaxies across time, as opposed to measuring properties of different galaxies at fixed mass across time. For all of our comparisons to observations, we explore all gas whereas observers typically measure abundances in HII regions specifically. However, Hernandez et al. (2021) compared observations of ionized and neutral gas-phase abundance gradients in M83, finding gradients for neutral gas to be $-0.17\,\mathrm{dex\,kpc^{-1}}$ and gradients for ionized gas to be $-0.03\,\mathrm{dex\,kpc^{-1}}$. This might imply that our measured gradients are much flatter than one would expect, given observations. Hernandez et al. (2021) did the same analysis excluding the nuclear region of M83 and found the neutral gas to be in much better agreement with ionized gas (a gradient of $-0.02\,\mathrm{dex\,kpc^{-1}}$. Additionally, we compare to observations which have measurements in broadly similar physical regions to those we analyze in the simulations, but they are not exactly the same.

We compared against observations of radial gradients in the MW, M31, and similar-mass galaxies at $z = 0$, finding broad agreement. We also connect our evolutionary trends with high-redshift observations of gas-phase abundances. In particular, we find that our MW/M31-mass galaxies all have negative radial gradients at $z \sim 0$ but had nearly flat radial gradients at $z \gtrsim 1$ (where the average stellar mass of the hosts is $M_{90} \approx 1.74 \times 10^{10}\mathrm{M}_{\odot}$). This trend agrees well with many observations of comparable mass galaxies at these higher redshifts (e.g. Queyrel et al., 2012; Stott et al., 2014; Wuyts et al., 2016; Patrício et al., 2019; Curti et al., 2020). However, some observational works have found strong negative radial gradients at these masses and redshifts (Wuyts et al., 2016; Carton et al., 2018; Wang et al.,

2020). Furthermore, while less common than negative radial gradients, some observations find some positive radial gradients at these redshifts as well (Queyrel et al., 2012; Wuyts et al., 2016; Carton et al., 2018; Wang et al., 2020) which we do not find in any of our galaxies. In general, we find that the steepening of radial gradients with time in our simulations is consistent with observational results and follows the quantitative trends in other theoretical analyses, both in simulations (e.g Ma et al., 2017a) and in metallicity-evolution modeling (Sharda et al., 2021).

One of the most important aspects of our analysis is quantifying azimuthal variations in gas abundances and comparing their strength relative to radial gradients across cosmic time. With the advent of integral field spectroscopy, observations have begun characterizing 2-D abundance distributions in nearby galaxies. These works (Sánchez et al., 2015; Vogt et al., 2017; Ho et al., 2017, 2018) all find non-trivial azimuthal variations in nearby galaxies, for example, Kreckel et al. (2019) found variations of ≈ 0.05 dex at fixed radius, which agrees well with our results. However, some observations (e.g. Zinchenko et al., 2016) found no evidence for large-scale azimuthal variation in nearby galaxies. One of our key results/predictions is the evolution of azimuthal variations, which we predict were stronger at higher redshifts. Observations of gravitationally lensed systems now allow sub-kpc measurements at high redshift (Jones et al., 2013, 2015), making it possible to test this predicted evolution in more detail.

Kreckel et al. (2020) examined azimuthal variations in gas-phase [O/H] across eight nearby galaxies using PHANGS-MUSE optical integral field spectroscopy. While our technique for measuring azimuthal variations are not exactly comparable to their methods, we find similar results. In our analysis we focus on scatter in all gas by measuring a mean scatter in angular bins of varying size, so we in effect measure the azimuthal inhomogeneities of random patches of gas at a given radius. In contrast to this, Kreckel et al. (2020) measure abundances specifically in HII regions and determine scatter by first subtracting off the

radial gradient and then centering apertures of various sizes on individual HII regions and measuring the scatter between the HII regions contained within the aperture. They find a slight scale dependence associated with the scatter, which we also see at $z = 0$, with the scatter on scales larger than $\approx 3\,\mathrm{kpc}$ being $\approx 0.05\,\mathrm{dex}$. The small-scale scatter in Kreckel et al. (2020) ($\approx 0.02\,\mathrm{dex}$) is slightly smaller than the $z = 0$ scatter we observe, but this could be attributed to the discrepancy in our methods. HII regions are likely better mixed in abundances than random patches of gas, so our analysis may be artificially inflating the typical azimuthal scatter of the gas from which stars are forming. However, centering on HII regions is beyond the scope of our analysis, and in future work we will examine azimuthal variations in newly formed stars, which may be closer to the values in HII regions.

Krumholz & Ting (2018) derived the expected correlation function of metal distribution in galaxies across space and time using a stochastic diffusion model. While we did not explore the correlation function of metals, we did examine homogeneity as a function of azimuthal scale, which we can compare broadly with their work. They found that gas-phase abundances produced primarily through core-collapse supernovae, in MW-like conditions near the solar circle, are correlated on scales of $\approx 0.5 - 1\,\mathrm{kpc}$ giving an expected scatter of $0.04 - 0.1\,\mathrm{dex}$. From fully cosmological simulations our results for azimuthal scatter on scales of $\approx 1\,\mathrm{kpc}$ near the solar cylinder agree well with their predicted range.

All of our results agree with Ma et al. (2017a), who analyzed radial gradients of abundances in the FIRE-1 simulations across a much wider galaxy mass range. In comparing with other theoretical/simulation works, our gradients in [O/H] at $z = 0$ fall between the gradients Hemler et al. (2021) measured in the TNG50 simulations ($\approx -0.02\,\mathrm{dex\,kpc^{-1}}$) and the gradients Gibson et al. (2013) measured in the MaGICC and MUGS simulations ($\approx -0.04\,\mathrm{dex\,kpc^{-1}}$). In particular, Hemler et al. (2021) found a gradual flattening of the gradients with time, which could come from an 'inside-out' growth of galaxies wherein star formation, hence elemental enrichment, proceeds from the inner galaxy to the outer galaxy

(e.g. Prantzos & Boissier, 2000; Bird et al., 2013). The flat(ter) radial gradients at earlier times in our galaxies result from higher turbulence and outflows that frequently eject much of the ISM at those times, perturbations such as mergers and rapid gas infall result in the velocity dispersion of gas particles dominating over their rotational velocity leading to galaxy-scale radial mixing (Ma et al., 2017a). As the disk settles over time, it becomes more rotationally supported, so stronger radial gradients can develop/persist. Our results qualitatively agree with those of the EAGLE simulations (Tissera et al., 2019), though as with TNG50, Tissera et al. (2019) found [O/H] gradients that are slightly shallower than ours, $\approx -0.011\,\mathrm{dex\,kpc^{-1}}$ at $z = 0$.

The evolution of our gas-phase abundance gradients disagrees with Agertz et al. (2021), who analyzed the VINTERGATAN simulation of the m12i initial conditions, performed using the adaptive mesh refinement code RAMSES. They found that the gas-phase profile of [Fe/H] becomes shallower over time (their Fig. 7), compared with our steepening with time. One possible explanation is the difference in hydrodynamic solvers: we use the mesh-free finite-mass (MFM) quasi-Lagrangian method in Gizmo, coupled with explicit modeling of sub-grid mixing, while the AMR simulation of VINTERGATAN induces significantly more mixing in gas (complete mixing on the scale of an individual cell), which may contribute to the flattening of their gradient over time. However, as shown in Appendix. A.3, the qualitative steepening of the gradient we observe with time is independent of the strength of our diffusion coefficient.

Galactic evolution models often simplify the abundance distributions of gas in galaxies to a 1-D model (e.g. Minchev et al., 2018; Mollá et al., 2019a; Frankel et al., 2020), with azimuthal scatter assumed from measurements at $z = 0$. While this is a useful first step in understanding the abundance evolution of galaxies and testing chemical tagging, our results mean that this simplifying assumption overestimates the radial information content in elemental abundances, including how well chemical tagging can constrain the birth radius

of a star. On the one hand, the non-trivial azimuthal scatter that we find, especially at earlier times, complicates modeling the abundances of stars at a given radius. On the other hand, this likely makes individual GMCs more elementally distinct at fixed radius, providing greater discriminating power, as we will explore in future work. However, we do not explore the homogeneity of individual GMCs in this work. Recent work has started to pursue 2-D abundance evolution models (e.g. Mollá et al., 2019b) which may address this question, works such as Spitoni et al. (2019) find azimuthal abundance variations on the order of 0.1 dex, twice what we find at $z = 0$ in our simulation suite.

Related to our analysis of a transition epoch, after which radial variations dominate over azimuthal scatter as the disk settles, Yu et al. (2021) examine the transition epoch from 'bursty' to 'steady' star formation and disk settling in the same simulations. We checked that their measurement of this bursty/steady transition agrees moderately well with the transition epoch that we present here, at least at smaller radius (4 kpc). We find weaker agreement for our transition times at larger radii (8 and 12 kpc). Furthermore, our transition times are consistently earlier (~ 3 Gyr at $R = 4$ kpc and ~ 1 Gyr at $R = 12$ kpc) than those in Yu et al. (2021), with the transition times on average being more similar at larger radii. Thus, we find a broad correlation between the transition from bursty to smooth star formation and the transition from azimuthal to radial abundance variations, but with significant scatter and some time delay.

Our simulations show the importance of considering azimuthal variations in addition to radial variations when studying gas-phase elemental abundance distributions. This is particularly important in the context of chemical tagging; in order to accurately identify the birth locations of stars using elemental abundances the initial conditions of stars need to be well defined. As we showed in Section. 2.4.5 azimuthal variations in abundance are greater than or comparable to radial variations at earlier times, so chemical tagging models that only account for radial variations will fail to accurately capture the scatter in abundances at

a given radius. This could lead to incorrectly assigning stars as co-natal, or vice versa. We also fit the elemental distributions of our galaxies at different radii at different times, finding that they shift from negative to zero skewness over time. While skew-normal fits are not a perfect fit for the elemental distributions of our galaxies at all redshifts, they more accurately represent the distributions than a Gaussian. Thus, using our measured distributions would be a useful step in building more accurate abundance evolution models including for chemical tagging.

Next generation telescopes are crucial for testing the predictions of gas-phase abundance homogeneity presented in this work, particularly the predictions for azimuthal scatter at high redshifts. Current measurements of azimuthal scatter in abundances have been restricted to nearby galaxies. However, with the advent of JWST NIRSPec IFU and next-generation adaptive-optic IFUs on telescopes like IRIS and TMT, spatially resolved measurements of metallicities in distant galaxies are feasible, providing tests of our predictions for azimuthal scatter and the transition redshifts when it becomes sub-dominant.

This work is the first step of testing the limits of chemical tagging in the FIRE simulations. In the future we will examine the degree to which these results for gas are mirrored in newly formed stars across time. Combining those results with measurements of the dynamical evolution of stars in our simulations, we more directly will test the efficacy of chemical tagging of stars in the FIRE simulations.

Acknowledgements

We thank Francesco Belfiore for providing abundance gradient measurements for galaxies from the MaNGA survey. We also thank Tucker Jones, Ryan Sanders, Joss Bland-Hawthorn, Trey Wenger, and Dana Balser for useful discussions that improved this manuscript. We thank the anonymous reviewer for useful comments that improved this article.

MAB and AW received support from NASA through ATP grants 80NSSC18K1097 and 80NSSC20K0513; HST grants GO-14734, AR-15057, AR-15809, and GO-15902 from STScI;

a Scialog Award from the Heising-Simons Foundation; and a Hellman Fellowship. Support for SRL was provided by NASA through Hubble Fellowship grant #HST-JF2-51395.001-A awarded by STScI, which is operated by AURA, Inc., for NASA, under contract NAS5-26555. We ran simulations using XSEDE supported by NSF grant ACI-1548562, Blue Waters via allocation PRAC NSF.1713353 supported by the NSF, and NASA HEC Program through the NAS Division at Ames Research Center. CAFG was supported by NSF through grants AST-1715216 and CAREER award AST-1652522; by NASA through grant 17-ATP17-0067; by STScI through grant HST-AR-16124.001-A; and by a Cottrell Scholar Award and a Scialog Award from the Research Corporation for Science Advancement. RF acknowledges financial support from the Swiss National Science Foundation (grant no 194814). We performed this work in part at the Aspen Center for Physics, supported by NSF grant PHY-1607611, and at KITP, supported by NSF grant PHY-1748958.

Data Availability

Full simulation snapshots at $z = 0$ are available for m12i, m12f, and m12m at ananke.hub.yt. A public version of the GIZMO code is available at `http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html`. Additional data including simulation snapshots, initial conditions, and derived data products are available at https://fire.northwestern.edu/data/. Some of the python code used to analyze these data includes the publicly available packages `https://bitbucket.org/awetzel/gizmo_analysis` (Wetzel & Garrison-Kimmel, 2020) and `https://bitbucket.org/awetzel/utilities`.

CHAPTER 3

3D elemental abundances of stars at formation across the histories of Milky Way-mass galaxies in the FIRE simulations

Published as Matthew Bellardini, Andrew Wetzel, Sarah Loebman, and Jeremy Bailin in Monthly Notices of the Royal Astronomical Society, Volume 514, Issue 3, August 2022, Pages 4270-4289, `https://doi.org/10.1093/mnras/stac1637`

3.1. Abstract

We characterize the 3-D spatial variations of [Fe/H], [Mg/H], and [Mg/Fe] in stars at the time of their formation, across 11 simulated Milky Way (MW)- and M31-mass galaxies in the FIRE-2 simulations, to inform initial conditions for chemical tagging.The overall scatter in [Fe/H] within a galaxy decreased with time until $\approx 7\,$Gyr ago, after which it increased to today: this arises from a competition between a reduction of azimuthal scatter and a steepening of the radial gradient in abundance over time.The radial gradient is generally negative, and it steepened over time from an initially flat gradient $\gtrsim 12\,$Gyr ago.The strength of the present-day abundance gradient does not correlate with when the disk 'settled'; instead, it best correlates with the radial velocity dispersion within the galaxy.The strength of azimuthal variation is nearly independent of radius, and the 360 degree scatter decreased over time, from $\lesssim 0.17\,$dex at $t_{\rm lb} = 11.6\,$Gyr to $\sim 0.04\,$dex at present-day.Consequently, stars at $t_{\rm lb} \gtrsim 8\,$Gyr formed in a disk with primarily azimuthal scatter in abundances.All stars formed in a vertically homogeneous disk, Δ[Fe/H]$\leq 0.02\,$dex within $1\,$kpc of the galactic midplane, with the exception of the young stars in the inner $\approx 4\,$kpc at $z \sim 0$.These results generally agree with our previous analysis of gas-phase elemental abundances, which reinforces the

importance of cosmological disk evolution and azimuthal scatter in the context of stellar chemical tagging.We provide analytic fits to our results for use in chemical-tagging analyses.

3.2. Introduction

Accurate models to describe the formation of the MW are crucial for interpreting and guiding observations of it. Current observational surveys, such as GALactic Archaeology with Hermes (GALAH; De Silva et al., 2015; Buder et al., 2018), Gaia-ESO (Gilmore et al., 2012), the Large Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Cui et al., 2012), and the Apache Point Galactic Evolution Experiment (APOGEE; Majewski et al., 2017; Ahumada et al., 2020; Jönsson et al., 2020), have measured abundances for hundreds of thousands of stars, and future surveys, including the Sloan Digital Sky Survey V (SDSS-V; Kollmeier et al., 2017), 4-metre Multi-Object Spectrograph Telescope (4MOST; De Jong et al., 2019), the WHT Enhanced Area Velocity Explorer (WEAVE; Dalton et al., 2012), and the MaunaKea Spectroscopic Explorer (MSE; The MSE Science Team et al., 2019) will extend the number of spectroscopically observed stars to the millions. These data, combined with high-fidelity models of galactic elemental enrichment, can offer tremendous insight into the formation history of the MW via 'chemical-tagging' (Freeman & Bland-Hawthorn, 2002).

Chemical tagging is a technique that leverages the elemental abundances of stars as an invariant to connect present-day observations of stars with their birth locations and times. By contrast, stellar orbital parameters change with time from mergers, accretion, and other dynamical scattering processes (for example Sellwood & Binney, 2002; Brook et al., 2004; Roškar et al., 2008a; Schönrich & Binney, 2009; Loebman et al., 2011).

One can consider chemical tagging in two regimes. 'Strong' chemical tagging associates stars with their birth cluster (for example Price-Jones et al., 2020), while 'weak' chemical tagging associates stars with their general birth time and location within the galaxy (for example Wojno et al., 2016; Anders et al., 2017). Crucially, either form of chemical tagging

relies on assumptions about the evolution of the spatial distribution of elemental abundances within a galaxy.

In the case of strong chemical tagging, gas clouds from which stars form must be sufficiently internally homogeneous, and the elemental abundances of individual gas clouds must be sufficiently unique from one another. Observations of star clusters indicate this first assumption is valid (for example Ting et al., 2012; Bovy, 2016). The extent to which the second requirement is met is less certain: observations of the MW and external galaxies indicate radial and azimuthal variations (for example Sánchez-Menguiano et al., 2016; Mollá et al., 2019b; Wenger et al., 2019; Kreckel et al., 2020) that could represent sufficiently unique abundances of star clusters.

Weak chemical tagging relies on similar assumptions applied instead to larger regions of the disk, for example, different stars at a given radius have homogeneous abundances that are elementally distinct from stars at other radii. In the extreme limits, one could imagine an elementally homogeneous disk, or an extremely clumpy disk in which all star clusters have unique abundances. The former would provide no spatially discriminating power, whereas the latter would in principle provide complete birth information, but would require complex models.

Critically, chemical tagging techniques rely on accurately modeling the evolution of the spatial scale of elemental abundance homogeneity of stars at birth. This provides constraints on the precision with which one chemically can tag stars. Bland-Hawthorn et al. (2010) previously explored this with a toy model and showed that most star clusters with masses below $\sim 10^5$ are internally homogeneous, but more work is needed to address the local and global degree of elemental abundance homogeneity in the MW.

Many works have measured the present abundance variations of stars in the MW. These observations indicate that the MW stellar disk has a negative vertical gradient, with more meta-rich stars closer to the disk midplane (for example Cheng et al., 2012; Carrell et al.,

2012; Boeche et al., 2014; Hayden et al., 2014), although the exact magnitude of the gradient varies between observations and exhibits radial dependence (Hayden et al., 2014). Additionally, a multitude of observations show that the stellar disk of the MW has a negative radial gradient in abundances (for example Boeche et al., 2013, 2014; Anders et al., 2014; Mikolaitis et al., 2014; Donor et al., 2018, 2020). However, the measured radial gradient varies significantly between observations, as well as varying with distance from the galactic midplane (for example Boeche et al., 2014; Hayden et al., 2014; Mikolaitis et al., 2014). Furthermore, Wang et al. (2019) showed that the magnitude of radial and vertical gradients is sensitive to stellar age.

Understanding how these variations change across cosmic time is imperative for chemical tagging. An often assumed consequence of 'inside-out' galaxy formation (Matteucci & Francois, 1989; Bird et al., 2013) is that, at larger lookback times, stellar disk radial gradients were steeper, and they have flattened with decreasing lookback time. This comes (naively) from assuming the strength of the abundance gradient necessarily follows from the strength of the overall surface-density gradient. Observations of mono-abundance stellar populations in the MW typically find older stellar populations have shallower metallicity gradients (e.g. Anders et al., 2017; Vickers et al., 2021), which these authors attribute to radial redistribution processes flattening the gradients of the oldest populations. However, another possibility is that older stars formed when the MW had a shallower radial gradient, that is, the MW's gradient has steepened over time.

The abundances of stars at formation trace that of gas, and both theoretical models and observations suggest that gradients steepen with time, though with significant uncertainty (see Mollá et al., 2019a, and references therein). Analyzing the FIRE-1 and FIRE-2 cosmological simulations of MW-mass galaxies Ma et al. (2017a) and Bellardini et al. (2021, hereafter B21) found that gas-phase abundance gradients steepen with decreasing redshift. The more analytic model for the evolution of gas-phase metallicity gradients by Sharda

et al. (2021) also indicates gradients tend to steepen with decreasing redshift, however, they may flatten between redshift ~ 0.2 and redshift 0. Additionally, high-redshift observations indicate gas-phase abundance gradients steepen with decreasing redshift (e.g. Curti et al., 2020).

In addition, several observations indicate that young stars show azimuthal variations in abundances across a galaxy (e.g. Luck et al., 2006; Lemasle et al., 2008; Pedicelli et al., 2009). Looking at B-type stars within 500 pc of the sun, Nieva & Przybilla (2012) found a scatter of ≈ 0.05 dex in [O/H]. Recently, using APOGEE data, Ness et al. (2022) determined that the median scatter in stellar abundance at fixed radius and time is $0.01-0.15$ dex for abundances generated via supernovae. However, in general, the evolution of azimuthal scatter of stellar abundances in the MW and MW-mass galaxies is not well understood, because it requires spatially resolved measurements of stellar abundances at high redshifts of lower-mass galaxies that are analogous to a MW progenitor.

Recent simulation work has emphasized the existence of azimuthal variations at $z = 0$. Solar et al. (2020) studied young star particles in 106 disks from an EAGLE simulation (Ref-L025N0752, initial gas mass resolution of $2.26 \times 10^5 \mathrm{M}_\odot$) and found the azimuthal variations lead to a scatter of $\sim 0.12\,\mathrm{dex} R_{\mathrm{eff}}^{-1}$ in the [O/H] radial gradient. Other simulation work (e.g. Grand et al., 2016; Di Matteo, 2016) has shown that azimuthal variations can arise in older stars within spiral galaxies because of streaming motion along non-axisymmetric features like bars and spiral arms. However, as of yet, cosmological simulations have not characterized the evolution of these azimuthal variations, especially for stars at the time of their formation, to provide context for chemical tagging.

In this paper, we characterize and provide fits for the cosmic evolution of 3-D abundance patterns in newly formed stars as a function of lookback time, to inform the initial conditions of stellar abundance distributions (prior to any post-formation dynamical changes), to inform the precision with which chemically tagging can recover stellar birth location and time. This

Simulation	$M_{90}^{\rm star}$ [10^{10}M$_\odot$]	$R_{90}^{\rm star,all}$ [kpc]	$R_{90}^{\rm star,young}$ [kpc]	Ex-situ percent
m12m[1]	10.0	11.9	12.7	6.8
Romulus[2]	8.0	14.8	16.9	8.2
m12b[3]	7.3	9.2	11.6	4.7
m12f[4]	6.9	13.5	17.0	3.0
Thelma[3]	6.3	11.7	15.1	3.2
Romeo[3]	5.9	14.2	16.8	2.9
m12i[5]	5.3	10.1	12.7	2.2
m12c[3]	5.1	9.2	11.8	5.2
Remus[2]	4.0	12.4	16.2	5.7
Juliet[3]	3.3	9.5	16.0	2.8
Louise[3]	2.3	12.6	17.3	1.9
Mean	5.9	11.7	14.9	4.2

TABLE 3.1. Stellar properties at $z = 0$ of our FIRE-2 MW/M31-mass galaxies. The first column lists the name of each simulation, the second column lists the stellar mass of the disk (see B21). The third and fourth columns list the cylindrical R_{90}^{*} (the cylindrical radius which includes 90% of the stellar mass) calculated using all stars and young (age < 250 Myr) stars, respectively, within a spherical aperture of radius < 30 kpc at $z = 0$. The fifth column lists the percentage of ex-situ stars, which we define to be stars that formed beyond a spherical radius $r > 30$ kpc comoving. The publication that introduced each simulation is: Hopkins et al. (2018)[1], Garrison-Kimmel et al. (2019b)[2], Garrison-Kimmel et al. (2019a)[3], Garrison-Kimmel et al. (2017)[4], Wetzel et al. (2016)[5].

builds upon our previous analysis (B21), where we explored the evolution of the elemental abundance distribution of all gas as a proxy for newly formed stars.

3.3. Methods

3.3.1. FIRE-2 Simulations. We use two suites of cosmological zoom-in simulations from the Feedback In Realistic Environments (FIRE) project[1] (Hopkins et al., 2018). We use 5 MW/M31-mass galaxies from the *Latte* suite (introduced in Wetzel et al., 2016), which have halo masses $M_{\rm 200m} = 1 - 2 \times 10^{12}M_\odot$, where $M_{\rm 200m}$ refers to the total mass within the

[1]FIRE project web site: http://fire.northwestern.edu

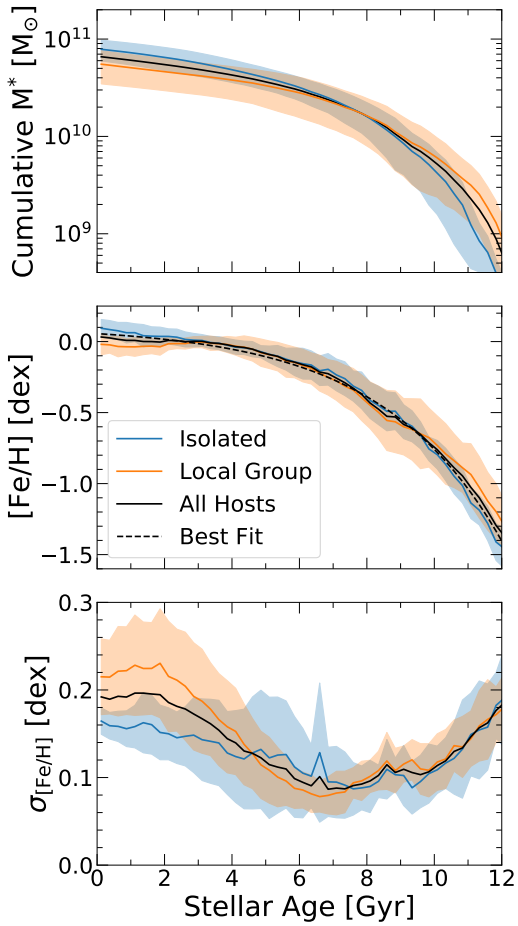


FIGURE 3.1. Formation histories of our 11 MW/M31-mass galaxies. We include all 'in-situ' stars that formed within a spherical radius $r < 30 \times a\,\mathrm{kpc}$ that remain within a geometrically defined disk (cylindrical radius $R < 20\,\mathrm{kpc}$ and vertical height $|Z| < 3\,\mathrm{kpc}$) at $z = 0$. We show the average across isolated galaxies (blue), LG-like galaxies (orange), and all galaxies (black). We also show the $1-\sigma$ scatter as shaded regions. **Top:** Cumulative stellar mass formed. While the isolated galaxies end up at slightly higher stellar mass at $z = 0$ (on average), the LG-like galaxies formed systematically earlier, as explored in Santistevan et al. (2020). **Middle:** Stellar [Fe/H], as a proxy for overall metallicity. The black dashed line shows our best fit (see Section 3.4.9). [Fe/H] increased until it saturated at $\approx 0.03\,\mathrm{dex}$, although the full galaxy sample spans $-0.1 - 0.20\,\mathrm{dex}$ at $z = 0$. The LG-like hosts experienced more rapid [Fe/H] enrichment at early times, with some reaching [Fe/H] $= -0.5 \approx 10\,\mathrm{Gyr}$ ago. **Bottom:** The scatter in [Fe/H] across the entire galaxy. The scatter at fixed age (time) was high for the oldest stars, then decreased over time down to a minimum $\approx 7\,\mathrm{Gyr}$ ago, after which it increased again to $z = 0$. Competition between decreasing bursty/clumpy star formation and increasing steepness of the radial gradient in [Fe/H] drive this shape.

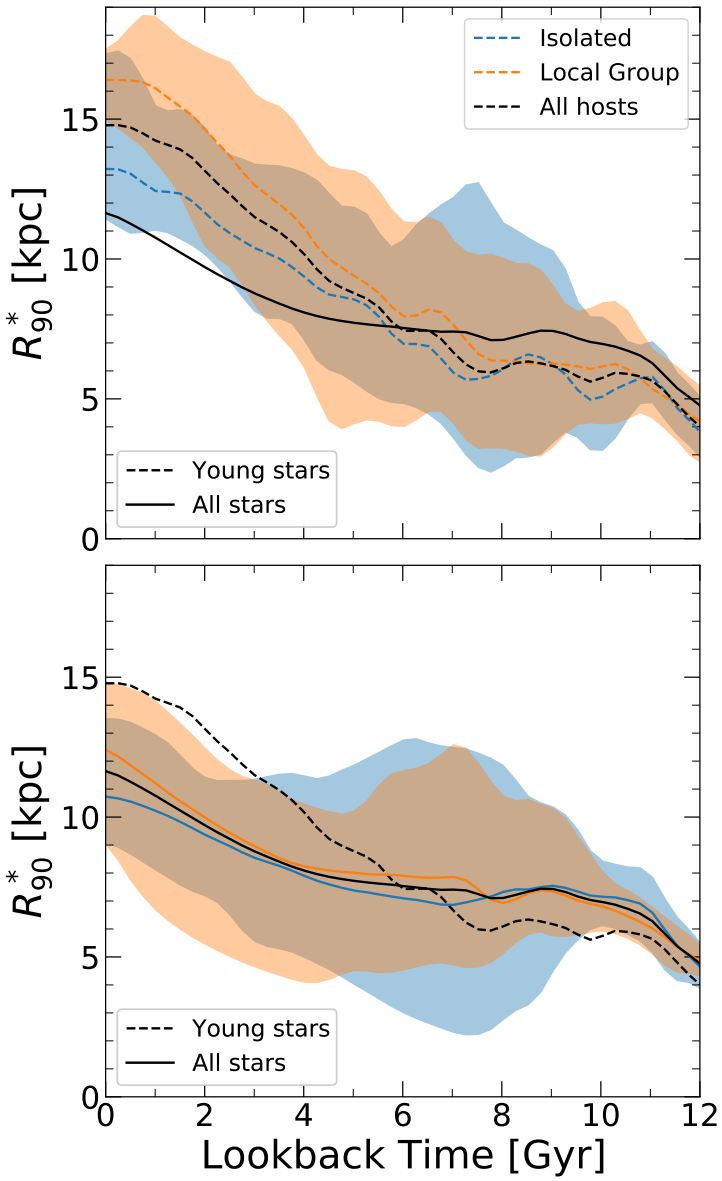


FIGURE 3.2. The stellar radius, which encloses 90% of the stars within $r < 30\,\mathrm{kpc}$ comoving (R^*_{90}), of the galaxies across time. We smooth all lines with a Gaussian filter with $\sigma = 250\,\mathrm{Myr}$. The dashed lines show R^*_{90} using only newly formed stars (ages $< 250\,\mathrm{Myr}$), and the solid lines show R^*_{90} for all stars. The lines show the mean R^*_{90} for isolated (blue), LG-like (orange), and all galaxies (black), with the shaded regions showing the range of the full distribution across our sample. R^*_{90} grew larger with time, reflecting inside-out radial growth. Prior to $\approx 7.5\,\mathrm{Gyr}$ ago, the size of the disks as determined by young versus all stars were similar. After that, R^*_{90} for young stars is larger than for all stars. Additionally, the sizes of disks of young stars in the LG-like galaxies are, on average, larger than those of isolated galaxies at late times.

radius within which the mean density is 200 times the mean matter density of the universe. The initial baryon particle mass in the simulations is $7070 M_\odot$ (however stellar mass loss leads to star particles having masses of $\approx 5000 M_\odot$ at $z = 0$), and the dark-matter mass resolution is $3.5 \times 10^5 M_\odot$. Star and dark-matter particles have fixed gravitational force softenings (comoving at $z > 9$ and physical at $z < 9$) with a Plummer equivalent of $\epsilon_{\rm star} = 4\,{\rm pc}$ and $\epsilon_{\rm dm} = 40\,{\rm pc}$. Gas cells have fully adaptive softening, which matches the hydrodynamic kernel smoothing, reaching a minimum softening length of 1 pc.

We also include 6 galaxies from the 'ELVIS on FIRE' suite of LG-like MW+M31 pairs (Garrison-Kimmel et al., 2019a,b). These have mass resolution $\sim 2\times$ better than *Latte*: the Thelma & Louise simulation has initial baryon particle masses of $4000 M_\odot$ and the Romeo & Juliet and Romulus & Remus simulations have initial baryon masses of $3500 M_\odot$.

All simulations use the FIRE-2 numerical implementations of star formation, stellar feedback, and fluid dynamics (Hopkins et al., 2018) using the Meshless Finite Mass hydrodynamics method of Gizmo (Hopkins, 2015). Gizmo conserves mass, energy, and momentum of particles to machine accuracy while enabling the adaptive hydrodynamic smoothing of gas elements based on their density.

The FIRE-2 model incorporates physically motivated models of star formation and stellar feedback. All simulations include the cosmic ultraviolet background from Faucher-Giguère et al. (2009). Gas cells experience metallicity-dependent radiative heating and cooling processes (across a temperature range of $10 - 10^{10}$K) including free-free, photoionization and recombination, Compton, photo-electric and dust collisional, cosmic ray, molecular, metal-line, and fine structure processes, accounting for 11 elements (H, He, C, N, O, Ne, Mg, Si, S, Ca, Fe). When discussing metallicities in this paper, we scale all elemental abundances to the solar values from Asplund et al. (2009).

Critical for our analysis, the simulations also model the sub-grid diffusion/mixing of elements in gas via turbulent eddies (Su et al., 2017; Escala et al., 2018; Hopkins et al., 2018).

B21 showed that the details of this implementation strongly affect small scale azimuthal abundance homogeneity. However, large scale azimuthal abundance variations, as well as vertical and radial trends are largely independent of the strength of our the sub-grid diffusion.

Stars form from gas that is self-gravitating, Jeans-unstable, cold ($T < 10^4$K), and molecular (following Krumholz & Gnedin, 2011). A newly formed star particle inherits the mass and elemental abundances of its progenitor gas cell. Each star particle represents a single stellar population, assuming a Kroupa (2001) stellar initial mass function, which evolves along standard stellar population models. We model time-resolved stellar feedback processes such as continuous mass loss from stellar winds, core-collapse and Ia supernovae, radiation pressure, photoionization, and photo-electric heating. We follow a combination of models (Van den Hoek & Groenewegen, 1997; Marigo, 2001; Izzard et al., 2004) synthesized in Wiersma et al. (2009) to model stellar winds and their yields. The rates of core-collapse and Ia supernovae come from STARBURST99 (Leitherer et al., 1999) and Mannucci et al. (2006) respectively. FIRE-2 nucleosynthetic yields follow Nomoto et al. (2006) for core-collapse and Iwamoto et al. (1999) for Ia supernovae.

We generated cosmological zoom-in initial conditions for all simulations embedded within cosmological boxes with side length $70.4 - 172\,\mathrm{Mpc}$, at $z \approx 99$ using MUSIC (Hahn & Abel, 2011). The simulations assume flat ΛCDM cosmology with parameters broadly consistent with the Planck Collaboration et al. (2020): $h = 0.68 - 0.71$, $\Omega_\Lambda = 0.69 - 0.734$, $\Omega_\mathrm{m} = 0.266 - 0.31$, $\Omega_\mathrm{b} = 0.0455 - 0.048$, $\sigma_8 = 0.801 - 0.82$ and $n_s = 0.961 - 0.97$. For each simulation we save 600 snapshots from $z = 99$ to $z = 0$, with typical time spacing of $\lesssim 25\,\mathrm{Myr}$.

We present the mass, size, and the ex-situ percentage of stars for all galaxies in our analysis in Table 3.1. We present the mass of all galaxies as measured in B21. We show our method for determining the size of the galaxies in Section 3.4.1. We define the ex-situ fraction as the fraction of stars currently within the galactic disk (defined geometrically with

cylindrical radius $R < 20\,\mathrm{kpc}$ and vertical height $|Z| < 3\,\mathrm{kpc}$) that formed outside of the spherical aperture of $r = 30\,\mathrm{kpc}$ comoving. We tested using a fixed spherical aperture of 30 kpc rather than a scale factor dependent aperture and found systematically larger disks at $t_{\mathrm{lb}} \gtrsim 6\,\mathrm{Gyr}$ largely driven by merger activity. We choose a scale factor dependent initial aperture to remove biases in R_{90} driven by starbursts in merging or close orbiting satellites.

3.3.2. Measuring stars at formation. We include only stars that end up within a geometrically defined disk at $z = 0$ ($R < 20\,\mathrm{kpc}$ and $|Z| < 3\,\mathrm{kpc}$). We present all results in terms of the properties of these stars at the time of their formation, to inform the 'initial conditions' for chemical tagging. Specifically, for each formation time, we look at stars in age bins of width 500 Myr and further subdivide these into 50 Myr age bins. We then analyze each 50 Myr age bin separately, and average the results of the 10 bins corresponding to the 500 Myr time window. This makes no difference to the radial and vertical trends we present, but is important for the azimuthal variations as, given the differential rotational dynamics of the disk, a 500 Myr age bin is larger than an orbital time.

To measure radial abundance profiles we first subdivide the stellar disk into annular bins of width 1 kpc and height $< 1\,\mathrm{kpc}$. For each 50 Myr age bin, within each annular bin, we store the mass-weighted mean stellar abundance. The total abundance profile is the mass-weighted mean of the profiles in each of 10 age bins making up the full 500 Myr age range.

To measure vertical abundance profiles, we first define annuli at different radii with width 2 kpc. We subdivide the annuli vertically into slices of height 100 pc. Using the absolute vertical height of star particles, we measure the mass-weighted mean profile in the same way as for the radial abundance profiles.

For the azimuthal abundance variations, we first define annuli at different radii with width 1 kpc and vertical height $< 1\,\mathrm{kpc}$. We analyze the azimuthal scatter at different scales. For various arclengths we subdivide each annulus into angular bins (the angular

bin sizes are defined such that 360° mod ϕ is zero, where ϕ is the angular size of the bin in degrees). The smallest arclength is no smaller than the 1 kpc width of the annulus and the largest arclength is the full annulus. For each 50 Myr age bin we measure the mass-weighted standard deviation of stellar abundances in each angular bin containing at least 3 star particles. For each size angular bin we report the average standard deviation as the mass-weighted mean of the standard deviation across all bins, for example, we average 10 bins at 360°, 20 bins at 180°, and so on.

We tested varying the minimum number of particles required per annular bin; for both 5 and 32 particles there is effectively no difference in the large-scale azimuthal scatter. However, reducing our minimum number of required particles enables us to measure smaller azimuthal scales. We also experimented with subdividing into different time intervals (25, 50, 100, 250, and 500 Myr intervals) and found that the azimuthal scatter for 50 Myr subdivisions was comparable to the azimuthal scatter for 25 Myr time intervals (the approximate time resolution of snapshots in the simulations) but provided better statistics. Increasing the time interval beyond 50 Myr increased the azimuthal scatter by up to a factor of ≈ 2 at large lookback times.

3.4. Results

3.4.1. Evolution of galaxy stellar mass, metallicity, and size. Fig. 3.1 shows the average formation histories of our 11 galaxies versus stellar age. We include stars within our geometrically defined disk at $z = 0$ that formed 'in-situ', that is, within spherical $r < 30 \times a$ kpc, where a is the scale factor at the time each star formed, of the main (most massive) progenitor.

Fig. 3.1 (top) shows the mean cumulative stellar mass of our 11 galaxies versus age. The blue line shows the isolated galaxies and the orange line shows the LG-like galaxies. The solid black line shows the mean across all galaxies. The shaded regions shows the $1 - \sigma$ scatter. The stellar mass increased from a mean of $5.5 \times 10^8 M_\odot$ 12 Gyr ago to a mean of

$6.6 \times 10^{10} M_\odot$ today. The average mass of the isolated galaxies is slightly ($\approx 1.4\times$) larger than that of the LG-like galaxies at $z = 0$. The LG-like galaxies show faster mass growth than the isolated galaxies at early times, in agreement with Garrison-Kimmel et al. (2019b) and Santistevan et al. (2020). At $t_{\rm lb} \gtrsim 8\,{\rm Gyr}$, the maximum mass difference between the LG-like hosts and the isolated hosts is $\approx 1.9 \times 10^9 M_\odot$.

Fig. 3.1 (middle) similarly shows the mean [Fe/H] of these galaxies for the same stellar selection. The dashed black line shows our best fit to the overall mean (see Section 3.4.9). The earlier mass assembly of the LG-like galaxies leads to their slightly higher metallicities at earlier times, which is even more pronounced in the upper boundary of the distribution (shaded orange region). At 10 Gyr ago, the mean [Fe/H] was ~ -0.75, with our most enriched (LG-like) galaxy reaching [Fe/H] $= -0.5$ already at that time. This agrees well with observations of old stars in the MW bulge (for example Bensby et al., 2017, and references therein) which find most dwarf stars with [Fe/H] $\lesssim -0.5$ are 10 Gyr or older. However, this metal enrichment largely saturates at late times: stars younger than $\approx 5\,{\rm Gyr}$ formed with a small range of [Fe/H] spanning approximately $-0.09 - 0.03\,{\rm dex}$. Overall, this indicates that the MW's LG environment may be key to understanding its early enrichment history.

Fig. 3.1 (bottom) shows the average scatter in [Fe/H] across the entire galaxy as a function of age. The scatter was larger at early times (that is, for the oldest stars today), and it decreased over time until $\approx 7\,{\rm Gyr}$ ago. However, after that the scatter increased over time, with the most recently forming stars having again high scatter. As we will explore below, this shape arose from the competition between two processes. The initial decrease in scatter at early time resulted from the decrease in azimuthal scatter as burstiness of star formation and turbulence in the ISM decreased. Conversely, the increase in scatter at late times arose from the steepening of the radial gradient. The epoch of minimum scatter in [Fe/H] ($\approx 7\,{\rm Gyr}$ ago) therefore coincided with the time at which the radial gradient was equal

to the azimuthal scatter, that is, when the radial gradient started to overtake the azimuthal scatter as the dominant source of inhomogeneity across the galaxy (see Section 3.4.7).

To provide context for our results on radial gradients, Fig. 3.2 shows how the galaxy sizes change with time. At each time, we fit R^*_{90} simultaneously (iteratively) with Z^*_{90}: these two define the radius and height of a cylinder in which the cumulative stellar mass of stars is 90% of total mass of stars within a spherical aperture of 30 kpc comoving. The top panel shows the size of the disk using only newly formed stars, with ages < 250 Myr, while the bottom panel shows the size of the disk using all stars. Fig. 3.2 shows the mean of the isolated (blue), LG-like (orange), and of all (black) galaxies as well as the full distribution (shaded region) of R^*_{90} for the isolated and LG-like hosts.

At lookback times $\gtrsim 5.5$ Gyr ago, the disk size as defined by young stars is smaller than the disk size as defined by all stars. However, after this, the disk size as determined by all stars is smaller than that determined by newly formed stars. This reflects the inside-out radial growth of galaxies: star formation proceeds across larger radii over time (for example Bird et al., 2013).

We also note size fluctuations at early times, which are similar to the 'breathing mode' fluctuations driven by stellar feedback in low-mass galaxies at $z \sim 0$ (El-Badry et al., 2016); however, our smoothing of the sizes partially washes out these short-time trends. We also measured R^*_{90} within a fixed spherical aperture (not scaling with the expansion scale factor), and we found overall similar trends, but with substantial scatter and strong fluctuations at early times, induced by mergers. Additionally, we tested the robustness of these results to varying the selection region and found no significant difference using a spherical aperture at $z = 0$ rather than a cylindrical 'disk-like' region.

For both size metrics, the LG-like galaxies have a larger R^*_{90} than the isolated galaxies at $z \sim 0$. The difference is larger for young stars, with the LG-like galaxies having disk sizes ~ 3 kpc larger than the isolated galaxies (~ 1 kpc larger when measuring all stars).

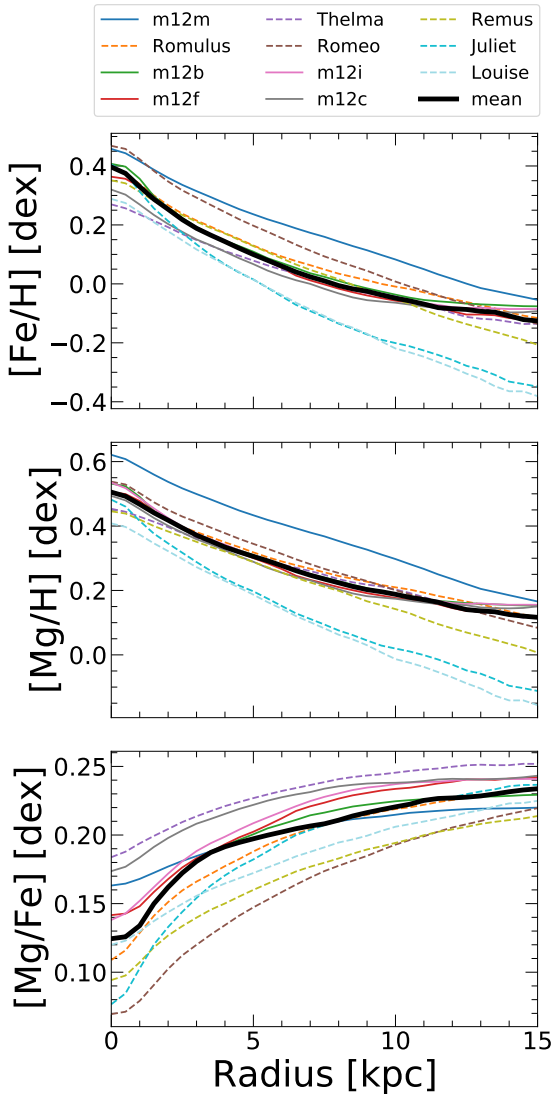


FIGURE 3.3. Radial profiles of elemental abundances of stars at formation, for stars that formed within the last 0.5 Gyr. We include all stars within a vertical height ± 1 kpc of the disk plane. We list the galaxies by decreasing stellar mass. The normalization of the [Fe/H] and [Mg/H] profiles scales roughly with stellar mass. The black line shows the mean, the solid lines show the isolated galaxies, and the dashed lines show the LG-like galaxies; we find no systematic differences between the LG-like and isolated galaxies. [Fe/H] and [Mg/H] both decrease with radius. The decrease is steeper in the inner region than in the outer region, with the typical transition occurring at $R \approx 5.5$ kpc. However, [Mg/Fe] increases with increasing radius, indicating more enrichment from Ia supernovae in the (older) inner disk than in the (younger) outer disk. This in turn helps to explain the origin of the gradient for young stars as well as the break in the abundance profile at $R \approx 5.5$ kpc.

This difference in R^*_{90} is consistent with the analysis of a subset of these galaxies (some at lower resolution, initial baryon mass of $2.8 - 3.2 \times 10^4 M_\odot$) in Garrison-Kimmel et al. (2018). We extend this analysis to 6 LG-like galaxies, all at full resolution (initial baryon mass of $3.5 - 4 \times 10^3 M_\odot$). This systematic difference implies that some aspect of the LG environment causes stars to form across a larger radius after the onset of disk formation (within the last $\sim 8\,\mathrm{Gyr}$), that is, causes more extended disk sizes. As Garrison-Kimmel et al. (2018) discussed, this may relate to stronger gas torques in LG-like environments, and/or this may relate to the earlier formation times of galaxies and their halos in LG-like environments, as Santistevan et al. (2020) showed. We defer a more detailed investigation to future work.

3.4.2. Radial abundance profiles at present day. We now show results for [Fe/H], [Mg/H], and [Mg/Fe]. [Fe/H] is an easily measured stellar abundance; it is sourced roughly equally by Type Ia supernovae and core-collapse supernovae, however, it is the primary metal produced in Ia supernovae. [Mg/H] is our representative α element, because it is the most 'pure' α element in the FIRE model, sourced almost entirely via core-collapse supernovae. [Mg/Fe] therefore reflects the relative enrichment from core-collapse versus Ia supernovae. We measure Mg as a representative α element rather than O (as in B21), because it is much easier to measure in stellar atmospheres than O. Thus, comparison of our results with observations is more straightforward.

Fig. 3.3 (top two panels) shows the radial profiles of [Fe/H] and [Mg/H] for stars younger than 500 Myr. Similar to our results for gas-phase abundances (B21), the galaxy-to-galaxy scatter in normalization primarily reflects the stellar mass-metallicity relationship (for example Tremonti et al., 2004; Ma et al., 2016). These profiles decrease monotonically with radius, because newly formed stars reflect the abundance of the gas, which also decreases with radius, both in these simulations (B21) and in observations (for example Mollá et al., 2019a, and references therein). The decrease in [Fe/H] across $0 - 15\,\mathrm{kpc}$ is $\approx 0.52\,\mathrm{dex}$,

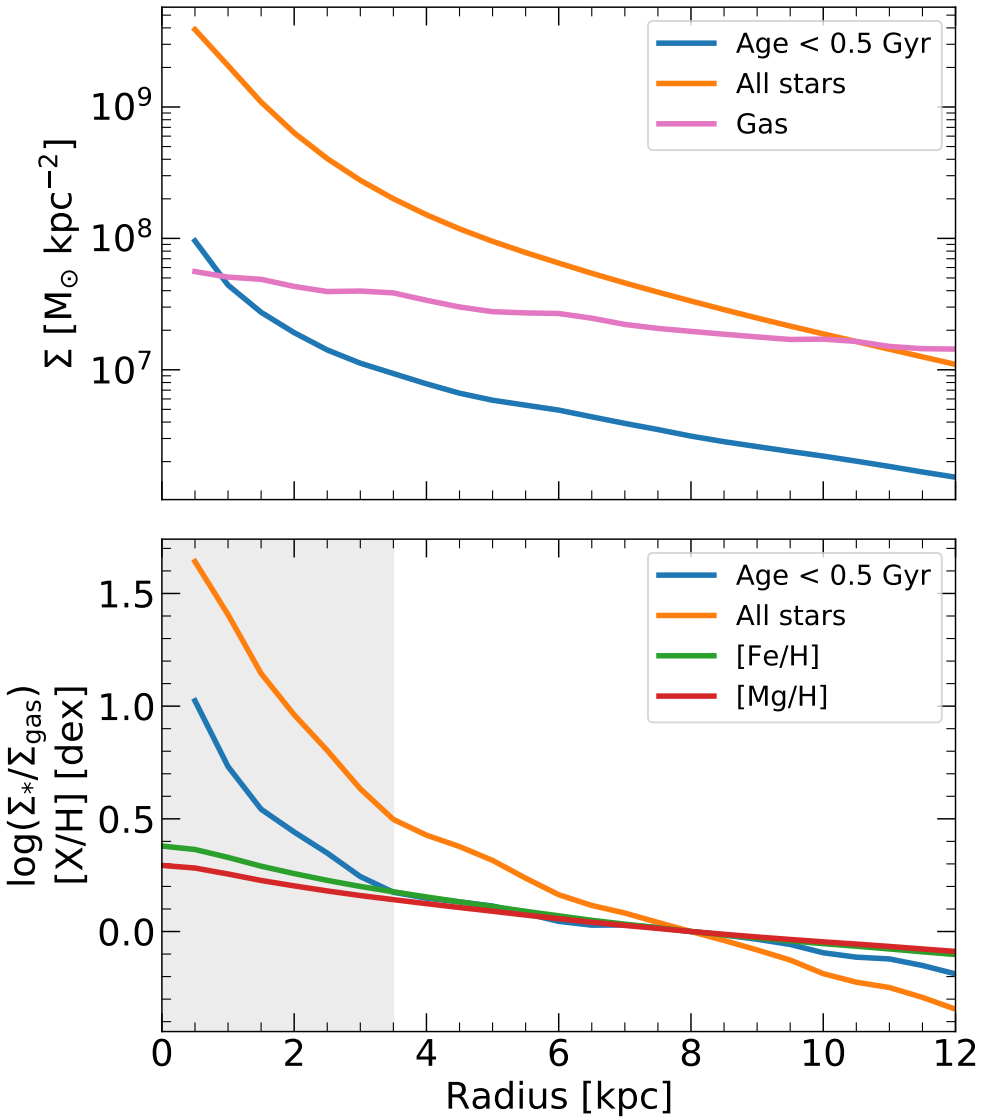


FIGURE 3.4. **Top**: Surface density versus radius, for all stars (orange), all gas (pink), and stars younger than 1 Gyr (blue), averaged over our 11 galaxies. The stellar surface densities exhibit steeper inner profiles (in their bulge-like regions) and shallower exponential profiles at larger radii. The gas profile is a shallower exponential at all radii, so gas dominates over stars at $R \gtrsim 10$ kpc. **Bottom**: The log ratio of stellar to gas surface density, and the [Fe/H] and [Mg/H] radial abundance, versus radius, all normalized to $R = 8$ kpc. The ratio of young stars to gas (approximately) should dictate the shape of the abundance profile for newly formed stars. We find reasonable agreement in the profile shapes at $R \gtrsim 3.5$ kpc. However, the surface density ratio rises more rapidly than the abundances in the inner few kpc (shaded gray), which may arise from metals being lost to outflows and/or dynamical redistribution of young stars.

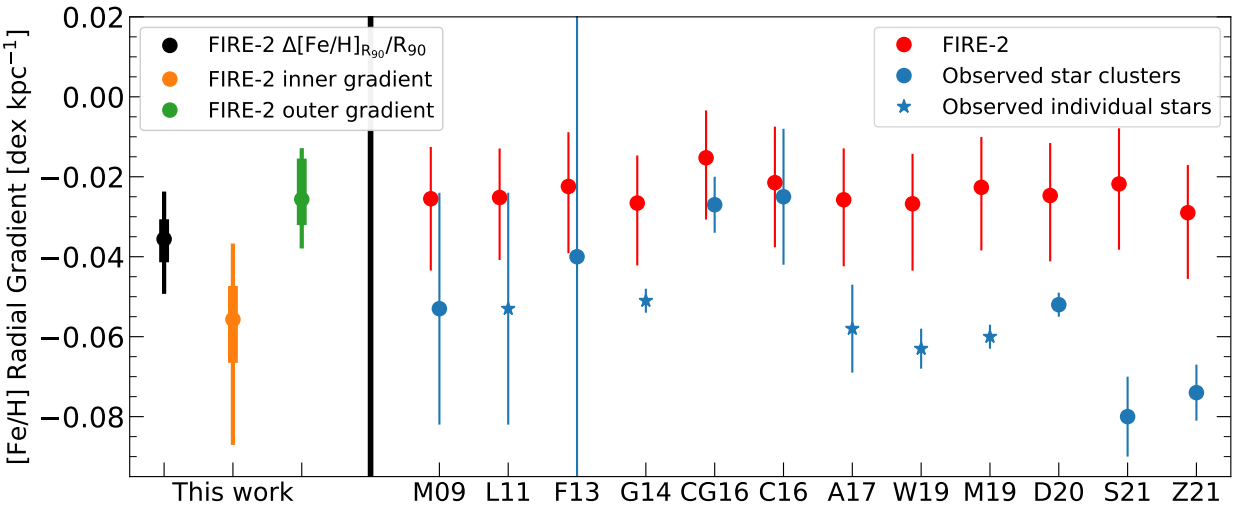


FIGURE 3.5. Radial gradients in [Fe/H] for young (ages $< 0.5\,\mathrm{Gyr}$) stars at $z = 0$, in the simulations and observed in the Milky Way (MW). For the simulations, our fiducial gradient is a total gradient (black), defined as the difference in metallicity between stars at R^*_{90} and at $R = 0\,\mathrm{kpc}$ divided by R^*_{90}. We also show the inner (orange) and outer (green) gradients of the stars measured by fitting a piecewise linear function to the radial profile at $R \leq R^*_{90}$. The thick lines show the 68th percentile of the simulations and the thin lines show the full distribution. We also compare best-fit linear profiles (red) to observational data (blue) of the MW from Magrini et al. (2009, M09), Luck & Lambert (2011, L11), Frinchaboy et al. (2013, F13), Genovali et al. (2014, G14), Cantat-Gaudin et al. (2016, CG16), Cunha et al. (2016, C16), Netopil et al. (2016, N16), Anders et al. (2017, A17), Wang et al. (2019, W19), Maciel & Andrievsky (2019, M19), Donor et al. (2020, D20), Spina et al. (2021, S21), and Zhang et al. (2021, Z21). Circular and star points show gradients determined via open clusters and individual stars, respectively. For each observational comparison, we fit the linear profiles in the simulations over the same radial range as observed. The radial gradients of our simulated galaxies are less steep than most observations of the MW across the same radial range. However, the inner gradients in our simulations are more consistent with the MW. As we showed in B21, these same simulations are steeper in gas-phase abundance gradients than most nearby MW-mass galaxies.

nearly identical to the $\sim 0.55\,\mathrm{dex}$ decrease in gas (B21). The decrease in [Mg/H] is weaker, at $\approx 0.39\,\mathrm{dex}$. The LG-like galaxies (dashed lines) have marginally steeper gradients in their outer disks. We explore potential causes of this in Section. 3.4.8 and defer a deeper investigation to future work.

Fig. 3.3 (bottom) shows that [Mg/Fe] increases with radius. Core-collapse supernovae source Mg and Fe relatively equal amounts so decreases in [Mg/Fe] indicate excess enrichment from type Ia supernovae, which produce more Fe than Mg. The low [Mg/Fe] in the inner galaxy likely reflects the inside-out radial growth of galaxies (for example Bird et al., 2021) seen in Fig. 3.2. Because stars in the inner galaxy are older on average, the inner galaxy has experienced more enrichment from Ia supernovae, while the outer disk is preferentially more enriched in Mg from core-collapse supernovae. This also explains why the [Fe/H] gradient is steeper than that of [Mg/H].

Fig. 3.3 also shows that the gradient is not linear: the slope is typically steeper in the inner region. We measure a break radius for each galaxy by fitting a two-component piecewise linear function to each profile. See Fig. B.1 for the break radius of each galaxy.

We explore the surface-density profiles in the galaxies to understand better the driver of the break in the abundance profile. Fig. 3.4 (top) shows the mean surface density profiles for stars younger than 1 Gyr (blue), all stars (orange), and all gas (pink) within 1 kpc of the galactic midplane, at $z = 0$. Fig. 3.4 (bottom) shows the log of the mean ratio of the stellar to gas surface density for young stars (blue) and all stars (orange), as well as the radial profiles of [Fe/H] (green) and [Mg/H] (red), all normalized at $R = 8$ kpc.

The elemental abundance in gas, and therefore in newly formed stars, approximately should scale with the ratio of stellar to gas mass, at least in the limiting case of instantaneous local enrichment. This is not the case in the inner ~ 3.5 kpc (shaded gray) of the galaxies, which implies metal loss from the inner galaxy. The details of this are beyond the scope of this paper, in which we emphasize the disk component.

The radial change beyond ~ 3.5 kpc is steeper for all stars than for young stars. This likely explains the shallower radial gradient seen in [Mg/H] relative to [Fe/H]. Mg is an α element and is primarily sourced by young stars in core-collapse supernovae, so the gradient in [Mg/H] traces the young stellar mass fraction. Fe is sourced roughly equally through

core-collapse and Ia supernovae, so [Fe/H] traces the stellar mass fraction of older stars on average. Thus a steeper gradient is expected for [Fe/H] than for [Mg/H].

The slope of the ratio of young-star to gas surface density is similar to that of the radial abundance profiles beyond $\sim 3.5\,\mathrm{kpc}$. This similarity indicates that this ratio, to first order, determines the abundance profile. Thus, this ratio likely partially drives the breaks in the abundance profiles. However, on an individual galaxy level, the breaks in the abundance profiles and the breaks in the young-star to gas surface density profiles are not always in agreement (see Appendix B.1 for more discussion).

3.4.3. Present-day radial profile compared to observations. Fig. 3.5 (left side) shows the total radial gradient in [Fe/H] for young (age $< 0.5\,\mathrm{Gyr}$) stars across our 11 galaxies in black, which we define as: the difference in [Fe/H] between $R = 0\,\mathrm{kpc}$ and $R = R^*_{90}$ divided by R^*_{90}:

$$\frac{\Delta[\mathrm{Fe/H}]_{\mathrm{R}^*_{90}}}{R^*_{90}} = \frac{[\mathrm{Fe/H}](\mathrm{R}=0) - [\mathrm{Fe/H}](\mathrm{R}=\mathrm{R}^*_{90})}{R^*_{90}} \tag{3.1}$$

Unlike all other results in this paper, here we use the locations of stars at $z = 0$, rather than their formation locations, to compare with observations (notably this changes the gradients by $\lesssim 0.002\,\mathrm{dex\,kpc^{-1}}$ on average). The thick line shows the $1-\sigma$ scatter and the thin line shows the full distribution. The median [Fe/H] gradient is $-0.036\,\mathrm{dex\,kpc^{-1}}$ with the full range of gradients spanning -0.049 to $-0.024\,\mathrm{dex\,kpc^{-1}}$.

Fig. 3.5 (left side) also shows the median, $1-\sigma$ scatter, and full distribution of the inner (orange) and outer (green) gradients. We compute these via a linear fit to each profile across $0 - R_{\mathrm{break}}$ and $R_{\mathrm{break}} - R^*_{90}$, where R_{break} is a free parameter of the fit. While a two-component piecewise linear function does not fully capture the shape of the abundance profile in all cases, we choose this functional form motivated by observations that find a break in the abundance radial profile (for example Andrievsky et al., 2004; Sestito et al., 2008; Magrini et al., 2009; Pancino et al., 2010; Frinchaboy et al., 2013; Hayden et al., 2014;

Korotin et al., 2014; Maciel & Andrievsky, 2019; Zhang et al., 2021). That said, the majority of observations that find such a break do not separate stars (or star clusters) by age, so the observed breaks might not simply reflect the behavior of stars at formation, but also could be affected by radial redistribution (for example Anders et al., 2017; Minchev et al., 2018; Quillen et al., 2018).

In our simulations, the gradient is steeper in the inner disk than in the outer disk. This is in contrast to the observational results of Maciel & Andrievsky (2019), that the abundance gradients of Cepheids tend to be steeper in the outer disk. However, our results agree with Magrini et al. (2009); Korotin et al. (2014), that the abundance gradients of Cepheids are flattest in the outer disk. We also agree with Eilers et al. (2022), who find that the MW's radial gradient in abundance is steeper in the inner disk than the outer disk for RGB stars. The inner [Fe/H] gradients span a range of -0.087 to $-0.037\,\mathrm{dex\,kpc^{-1}}$ with a median of $-0.056\,\mathrm{dex\,kpc^{-1}}$ and the outer gradients span a range of -0.038 to $-0.013\,\mathrm{dex\,kpc^{-1}}$ with a median of $-0.026\,\mathrm{dex\,kpc^{-1}}$.

Fig. 3.5 (right side) compares the [Fe/H] gradients for young stars in our simulations (red) and observations (blue) of individual stars (star points) and young star clusters (circles). In each case, we measure the gradient within a vertical height $< 1\,\mathrm{kpc}$ that spans the same radial range as each observation. The error bars span the full range across our 11 galaxies.

The [Fe/H] radial gradients of the young stars in our simulations are less steep than most radial gradients observed in the MW, especially for more recent observations. This agrees with our previous results on gas-phase abundance gradients in those simulations compared with the MW in B21. However, we also found that our gas-phase gradients agree well with M31, and they tend to be steeper than gradients observed in most nearby MW-mass galaxies. At face value, this suggests that the MW's abundance gradient is unusually steep compared with similar-mass galaxies (for example Boardman et al., 2020).

Observational distance uncertainties could affect these measurements of the MW: Donor et al. (2018) found that using different distance catalogs can change observed radial gradients by up to 40%. Another possibility that one might consider is that the MW disk settled unusually early compared to nearby galaxies and our simulations. However, as we show in Section 3.4.8, earlier disk settling in our simulations does not correlate significantly with a stronger gradient at $z = 0$. We do not think that our model for sub-grid diffusion of metals imparts shallowness in our radial gradients, because in B21 we showed that using a lower diffusion coefficient has little to no impact on our radial gradients. However, if our stellar feedback is too strong, this could drive excess turbulence in the gas, flattening our radial gradients.

In summary, our simulated radial gradients of [Fe/H] tend to be shallower than the MW when measured over the same radial range, though our distribution of inner and outer gradients does encompass the full range of observed MW radial gradients. In other words, we at least recover observed MW gradients within different regions of our galaxies, so any discrepancy may be simply in the location of this radial break. We also emphasize from B21 that our gradients are similar to M31 and similar to or steeper than nearby MW-mass galaxies. In future work we will compare our gradients to the MW across the full range of stellar ages.

3.4.4. Evolution of the radial abundance profile. Fig. 3.6 shows the radial profiles of abundances in newly formed stars at different lookback times. The solid line shows the mean profile at each lookback time out to the average R^*_{90}, the dashed line shows the profile beyond R^*_{90}, and the shaded region shows the $1 - \sigma$ scatter. Scatter in these galaxies' formation histories, combined with the mass-metallicity relationship, leads to different normalizations of abundances at different times, as in Fig. 3.1, which blurs the trends in the profiles, leading to scatter that primarily reflects different normalizations rather than

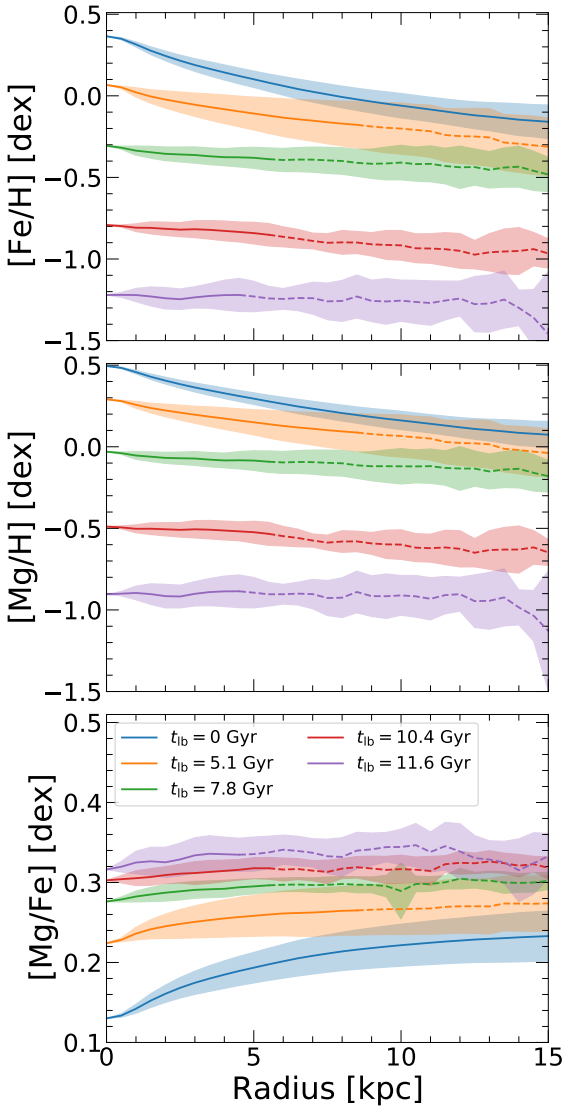


FIGURE 3.6. Radial profiles of elemental abundances of stars at formation across our 11 MW/M31-mass galaxies at various lookback times. Solid lines shows the profiles out to the average R^*_{90} at each lookback time, while the dashed lines show beyond that. The lines show the mean and the shaded regions show the $1-\sigma$ scatter across the galaxies. Because the galaxies had different stellar masses and different abundance normalizations at early times, we re-normalize to the mean abundance at $R = 0$ across the galaxies at each time. Similar to the gas-phase abundance profiles of these galaxies (see B21) the stellar abundance profiles are flattest at early times (oldest stars) and steepest at latest times (youngest stars). The combined evolution of the normalization and radial gradients leads to a degeneracy (for chemical tagging), such that stars born at different times and radii can have the same abundance, especially within the last ~ 8 Gyr.

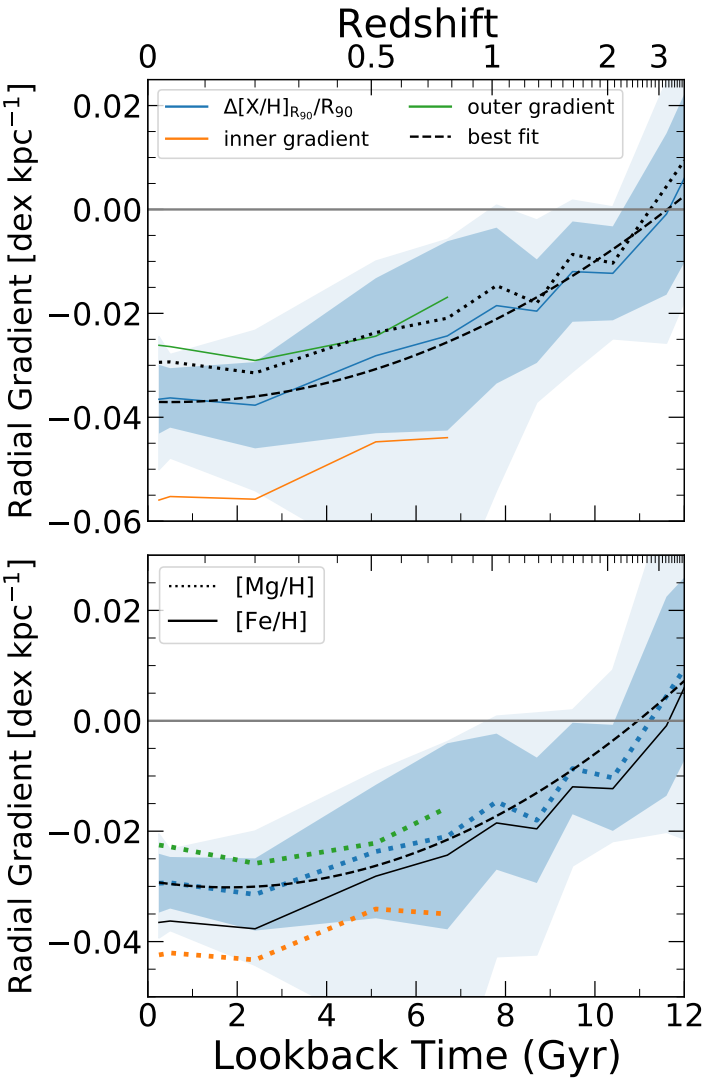


FIGURE 3.7. Radial gradients of elemental abundances of stars at formation versus lookback time. The blue line shows the overall gradient, defined as the difference between abundance at R^*_{90} and at $R = 0\,\mathrm{kpc}$ divided by R^*_{90}. The dark shaded region shows the $1-\sigma$ scatter, and the light shaded regions shows the full distribution across our 11 galaxies. The black dashed line shows our best fit to this evolution (see Section 3.4.9). The orange and green lines show the inner and outer gradients, respectively, measured by fitting a piece-wise linear function to the abundance profile (see Fig. 3.6), plotted at the look-back times where they reasonably fit the abundance profiles. The inner gradient is always steeper than the outer gradient. All gradients become steeper over time, and [Fe/H] is slightly steeper than [Mg/H]. The most extreme gradients in the full distribution reach $\sim -0.077\,\mathrm{dex\,kpc^{-1}}$ at $t_{\mathrm{lb}} = 6.7\,\mathrm{Gyr}$ and $\sim 0.032\,\mathrm{dex\,kpc^{-1}}$ at $t_{\mathrm{lb}} = 12\,\mathrm{Gyr}$ for [Fe/H] ($\sim -0.070\,\mathrm{dex\,kpc^{-1}}$ at $t_{\mathrm{lb}} = 6.7\,\mathrm{Gyr}$ and $0.039\,\mathrm{dex\,kpc^{-1}}$ at $t_{\mathrm{lb}} = 11.6\,\mathrm{Gyr}$for [Mg/H]).

the shapes of the profiles. Thus, we normalize the abundance profiles of all galaxies to the average abundance at $R = 0\,\mathrm{kpc}$ at each lookback time.

Fig. 3.6 (top two panels) shows that, as galaxies evolve, the average metallicity of newly formed stars increased at all radii. The increasing normalization agrees with the trends in Fig. 3.1. Subsequent generations of stars formed from gas that grew more enriched over time.

Beyond this normalization, the shapes of the radial gradients of [Fe/H] and [Mg/H] for young stars changed with time. In general the profiles got steeper over time, such that the profiles of stars that formed $\sim 12\,\mathrm{Gyr}$ ago were approximately flat, while the most recently formed stars have negative gradients. These radial profiles of stars at formation trace that of the gas, which B21 showed grew steeper over time in a similar way, a result of the inside-out radial growth of these galaxies (see Fig. 3.2).

Fig. 3.6 (bottom) shows that [Mg/Fe] tends to decrease for newly formed stars at all radii with increasing lookback time. [Mg/Fe] drops more in the inner galaxy than the outer galaxy, leading to a steepening positive profile over time. This qualitatively matches the results for gas-phase abundances in B21, likely because the older inner galaxy experienced more Ia supernovae, which preferentially enrich the gas (hence the newly forming stars) with Fe, whereas core-collapse supernovae preferentially enrich the younger outer galaxy with α elements like Mg.

Fig. 3.7 shows the evolution of radial gradients in [Fe/H] and [Mg/H] for newly formed stars versus lookback time. The blue line shows the mean of our fiducial radial gradient: $\Delta[\mathrm{Fe/H}]_{R^*_{90}}/R^*_{90}$. The dark blue shaded region shows the $1-\sigma$ scatter and the light blue shaded region shows the full distribution across our 11 galaxies. The orange and green lines show the inner and outer gradients via fitting a two-component piece-wise linear function (see Section 3.4.2). However, we only show the two-component gradients out to $\approx 6.7\,\mathrm{Gyr}$, because the functional form is not a good fit when the gradients are sufficiently flat. The

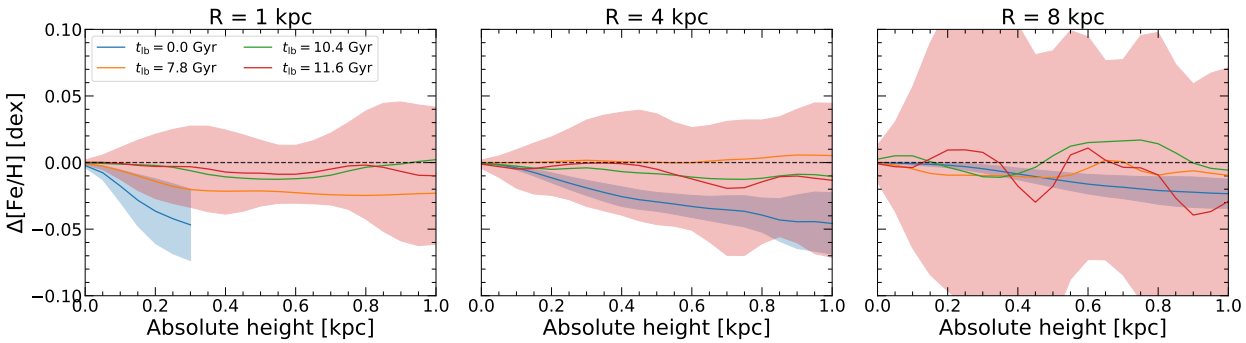


FIGURE 3.8. Vertical profile (relative to the disk mid-plane) of [Fe/H] of stars at formation, at 3 radial annuli centered on $R = 1\,\mathrm{kpc}$ (left) $R = 4\,\mathrm{kpc}$ (center) and $R = 8\,\mathrm{kpc}$ (right) with width $\pm 1\,\mathrm{kpc}$. The solid lines show the mean deviation in abundance at each height, and the shaded regions show the $1-\sigma$ scatter across our 11 galaxies. We smooth all lines with a Gaussian filter with $\sigma = 0.05\,\mathrm{kpc}$. [Fe/H] is generally invariant with increasing height from the midplane. The only exception is stars at present-day, which show a modest decrease in [Fe/H] and with increasing height ($\gtrsim 0.05\,\mathrm{dex\,kpc^{-1}}$ slope for all $R \lesssim 4\,\mathrm{kpc}$), tracing the abundance pattern of the star-forming gas in this region.

black dashed line in the top panel shows the best fit to the [Fe/H] evolution as we describe in Section 3.4.9.

The radial gradients became more negative with decreasing lookback time. The mean total gradient was $0.006\,\mathrm{dex\,kpc^{-1}}$ ($0.009\,\mathrm{dex\,kpc^{-1}}$) for [Fe/H] ([Mg/H]) for stars that formed 12 Gyr ago, steepening to $-0.037\,\mathrm{dex\,kpc^{-1}}$ ($-0.029\,\mathrm{dex\,kpc^{-1}}$) for [Fe/H] ([Mg/H]) for stars that formed < 0.5 Gyr ago. The trend of steeper radial gradients in [Fe/H] is consistent at nearly all lookback times. We discuss the functional form of this evolution in Section 3.4.9.

Fig. 3.7 also shows our two-component fit to the gradients, which we do not show at large lookback times, when the profiles were flat (see Fig. 3.6). However, characterizing the transition of the disk from a single-component to a multi-component profile is beyond the scope of this paper. Similar to our measurement of the total gradient, both the inner and outer gradients steepen over time. For the stars forming at $z \sim 0$, the mean inner gradient is $-0.056\,\mathrm{dex\,kpc^{-1}}$ ($-0.042\,\mathrm{dex\,kpc^{-1}}$) and the mean outer gradient is $-0.026\,\mathrm{dex\,kpc^{-1}}$ ($-0.023\,\mathrm{dex\,kpc^{-1}}$), for [Fe/H] ([Mg/H]). The inner gradient is consistently steeper than the

outer gradient at all times, which, as we argue from Fig. 3.4, indicates that the ratio of stellar to gas mass always had a steeper profile in the inner galaxy.

3.4.5. Vertical profile. Fig. 3.8 shows the mean change in abundance as versus distance from the galactic midplane for newly formed stars within 2 kpc wide bins centered at 3 radii, $R = 1$ kpc (left), $R = 4$ kpc (center), and 8 kpc (right), at various lookback times. We measure this change with respect to the midplane abundance value of each host. We smooth all profiles with a Gaussian filter (using scipy.ndimage.gaussian_filter1d with $\sigma = 0.05$ kpc). The solid lines show the mean, and the shaded regions show the $1 - \sigma$ scatter across our 11 galaxies. For clarity we show only the scatter at present day and at the largest lookback time.

The [Fe/H] and [Mg/H] (not shown) profiles show little to no significant variation with increasing distance from the midplane. The mean profiles show slight deviations up to 1 kpc, but any systematic trends are typically smaller than the galaxy-to-galaxy $1 - \sigma$ scatter, and/or the strength is typically smaller than the measurement precision of most stellar surveys. These flat vertical profiles are unsurprising, given the similar gas-phase results we explored in B21. We thus conclude that vertical gradients in abundances are generally negligible.

The one exception is young stars at present-day in the inner disk, which have a systematic modest negative vertical gradient ($\gtrsim 0.05\,\mathrm{dex\,kpc^{-1}}$ out to approximately 4 kpc). This declining metallicity with increasing distance from the midplane agrees with some observations of the MW (for example Katz et al., 2011; Hayden et al., 2014; Xiang et al., 2015; Wang et al., 2019). We find that this trend is caused by the star-forming gas, that is, it is not affected by any post-formation stellar dynamics on timescales $\lesssim 20$ Myr. In particular, unlike at larger radii, the metallicity distribution of star-forming gas at small radii near the midplane is highly non-Gaussian, indicating that enrichment timescale is shorter than the mixing timescale. In other words, given both the higher star-formation (metal enrichment)

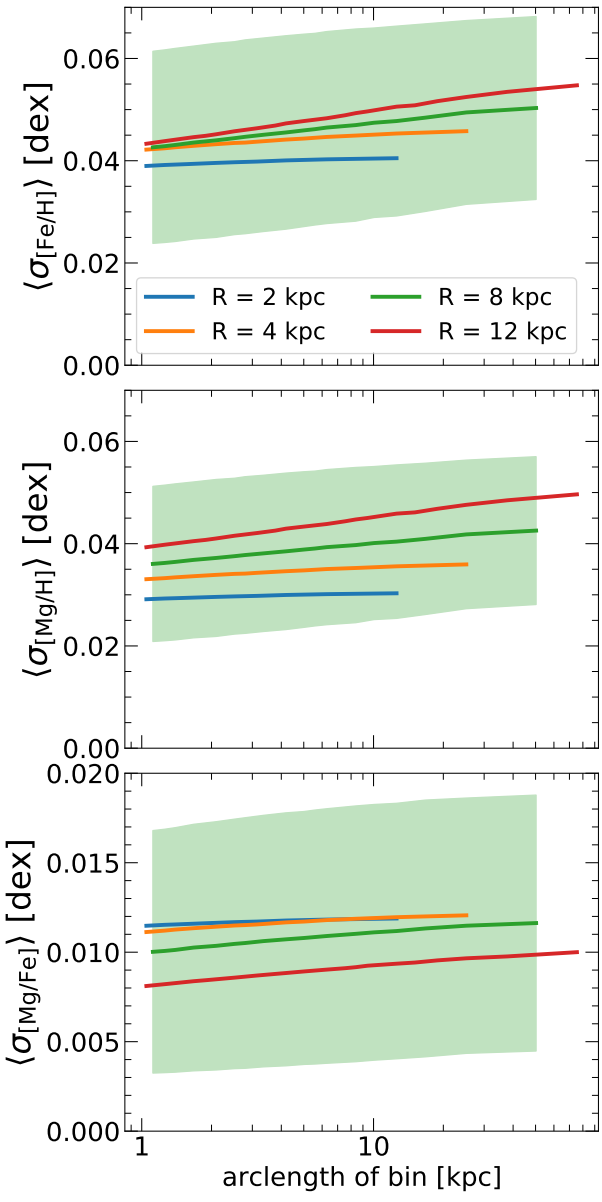
density and the stronger gravitational potential in the inner galaxy, gas is not able to mix as efficiently in the vertical direction, leading to enhanced enrichment in the midplane.

We additionally explored the vertical profile of [Mg/Fe] (not shown here). [Mg/Fe] increases with distance from the midplane in the inner disk at present-day. Here, both [Mg/H] and [Fe/H] decrease (slightly) with height, so the stars at larger heights are somewhat less enriched. The [Mg/H] gradient is slightly weaker, though, which may result from vertical enrichment being more affected by star-forming winds, enriched preferentially in Mg from core-collapse supernovae.

We caution that these simulations do not include black-hole feedback (see Wellons et al., 2023), which could affect these vertical gradients in the inner galaxy.

3.4.6. Azimuthal scatter. Fig. 3.9 shows the azimuthal scatter of elemental abundances at formation of stars at $z = 0$ with ages $< 0.5\,\mathrm{Gyr}$ at 4 radii ($R = 2, 4, 8, 12\,\mathrm{kpc}$) as a function of azimuthal arclength. The solid lines show the mean scatter, and the shaded region highlights the standard deviation for stars that formed at our fiducial solar radius of $R = 8\,\mathrm{kpc}$.

The azimuthal scatter at $z = 0$ increases modestly with radius and depends only weakly on azimuthal bin size. At $R = 2\,\mathrm{kpc}$ the mean scatter is essentially independent of the angular bin arclength and the scatter within the whole annulus is $\sim 0.01\,\mathrm{dex}$ smaller than the mean scatter in the annulus at $R = 8\,\mathrm{kpc}$. At $R = 8\,\mathrm{kpc}$ the scatter depends slightly on azimuthal bin width, increasing from $0.043\,\mathrm{dex}$ to $0.05\,\mathrm{dex}$ for [Fe/H] ($0.036\,\mathrm{dex}$ to $0.043\,\mathrm{dex}$ for [Mg/H]). Fig. 3.9 shows that the azimuthal scatter of [Mg/Fe] decreases slightly with increasing radius, in contrast with the trends for [Fe/H] and [Mg/H]. However, the radial dependence is smaller than the host-to-host scatter. Importantly, we do not center our bins on individual star-forming regions (star clusters), so, for all panels in Fig. 3.9, the scatter does not go to $0\,\mathrm{dex}$ at small arclengths, as one might expect.



f

FIGURE 3.9. Azimuthal scatter of elemental abundances versus arclength, for stars at formation within the last 500 Myr at $z = 0$, using annuli 1 kpc in width centered on 4 different radii, averaged across our 11 galaxies. The shaded region shows the $1-\sigma$ scatter in a fiducial solar cylinder ($R = 8$ kpc). The azimuthal scatter in [Fe/H] and [Mg/H] increases with arclength and with radius. The scatter in [Mg/Fe] decreases slightly with radius. The 360° scatter in [Fe/H] at $R = 8$ kpc is ≈ 0.05 dex, which agrees well with the scatter in gas in B21. The minimal dependence on azimuthal bin width (arclength) agrees well with the dependence at $z = 0$ in B21. We do not center these bins on star-forming regions, so even on scales ~ 1 kpc the scatter remains $\sim 0.03 - 0.04$ dex.

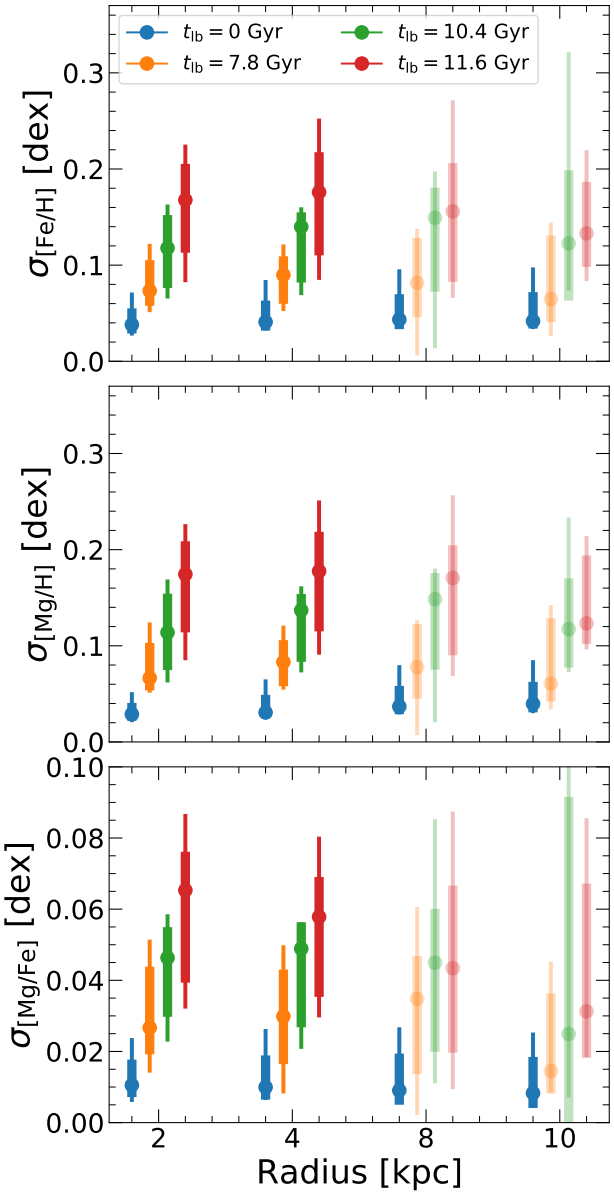


FIGURE 3.10. Full (360°) azimuthal scatter of newly formed stars as a function of radius at various lookback times. The points show the median scatter, the thick lines show the $1-\sigma$ scatter, and the thin lines show the full distribution across our 11 galaxies. The lightly shaded points show the scatter beyond the average R^*_{90} of the galaxies. The azimuthal scatter shows little to no dependence on radius. We show only the largest-scale (360°) azimuthal scatter, because as Fig. 3.9 shows, the scatter depends minimally on azimuthal bin width. By contrast, the azimuthal scatter increases with lookback time. The median scatter at 8 kpc increases from $\sim$ 0.04 dex to $\approx$ 0.16 dex from $t_{lb} = 0$ Gyr to 11.6 Gyr. However, this time dependence is weaker than for all gas in these galaxies (see B21 Fig.7).

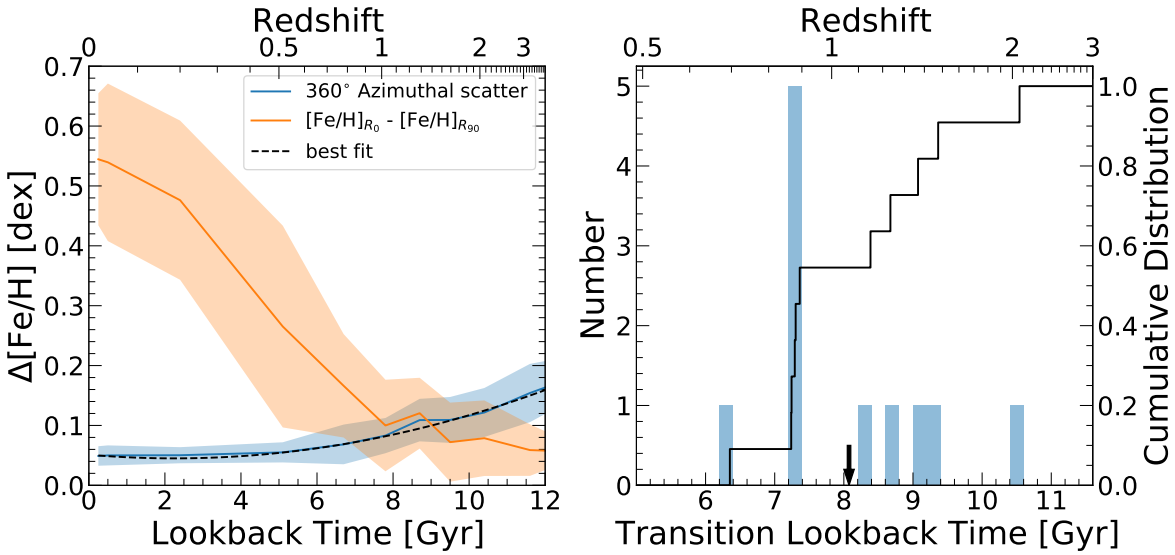


FIGURE 3.11. **Left:** The magnitude of radial versus azimuthal variations in [Fe/H] for newly formed stars, versus lookback time. The orange line shows the absolute change from $R = 0\,\mathrm{kpc}$ to R^*_{90}, while the blue line shows the 360° azimuthal scatter. The solid lines show the average and the shaded regions show $1-\sigma$ scatter across our 11 galaxies. The intersection of these 2 lines defines a transition age: stars older than this formed in a galaxy dominated by azimuthal scatter, while stars younger than this formed in a disk in which radial variations dominated. **Right:** A histogram of this transition lookback time across our 11 galaxies. The black line shows the cumulative distribution. The black arrow shows the average ($\approx 8\,\mathrm{Gyr}$ ago), which is $0.5-1\,\mathrm{Gyr}$ earlier than the average transition time for gas (B21).

Fig. 3.10 shows the evolution of the azimuthal scatter, specifically, showing the 360° scatter versus radius at various lookback times. We show only the 360° scatter here, because the difference between the scatter at the smallest and largest scales is minimal ($< 0.018\,\mathrm{dex}$) for all lookback times and radii. The points show the median scatter, the thick lines show the $1-\sigma$ scatter, and the thin lines show the full distribution across our 11 galaxies. The lightly shaded points indicate radii that are larger than the average R^*_{90} at a given lookback time.

The azimuthal scatter at all radii and all azimuthal bin sizes generally increase with increasing lookback time. Stars that have formed more recently formed in a more homogeneous azimuthal disk than stars that formed earlier. This agrees with the trends for gas

abundances in B21, though at large lookback times the scatter in newly formed stars is generally smaller than the scatter seen for all gas (see Appendix B.2). The smaller scatter in stars results in part from stars forming from gas that is preferentially metal rich, especially at large lookback times (for example B21) and at large radii. Additionally, newly formed stars are more spatially clustered than all gas cells. This explains the weaker (shallower) dependence on azimuthal bin size for young stars than for gas: the fraction of bins containing no mass is larger for stars than for gas, so increasing the binsize does not necessarily include more/different stars per bin.

The azimuthal scatter for stars shows little dependence on radius, in contrast to the azimuthal scatter for gas (B21). Here, the change in scatter with radius is in general less than ~ 0.02 dex. Thus, stars throughout the galaxy form in essentially equally azimuthally homogeneous conditions.

Fig. 3.10 (bottom) shows the evolution of azimuthal scatter for [Mg/Fe]. Similar to gas (B21), the scatter in [Mg/Fe] is much smaller at all radii than the scatter in [Fe/H] or [Mg/H]. However, the scatter in [Mg/Fe] follows the same trend of increasing with increasing lookback time. The scatter in [Mg/Fe] shows less dependence on azimuthal bin width than the scatter for [Fe/H] and [Mg/H] with a maximum change of ≈ 0.006 dex between the smallest and largest scales. The galaxy-to-galaxy standard deviation is also smaller for [Mg/Fe] than for [Fe/H] and [Mg/H].

3.4.7. Strength of azimuthal versus radial variation. Given the first-order approach of chemical evolution/chemical tagging models to neglect azimuthal scatter (for example Minchev et al., 2018; Mollá et al., 2019a; Frankel et al., 2020), it is critical to identify when the assumption of minimal azimuthal scatter is valid. These models are only accurate representations of the MW when azimuthal abundance variations are smaller than radial abundance variations. The general increase in the steepness of radial gradients with time (seen in Section. 3.4.4) and the general decrease in azimuthal scatter with time (seen in

Section. 3.4.6) implies there must be some transition time prior to which azimuthal scatter is the dominate source of abundance variations and after which radial variations dominate.

Fig. 3.11 (left) compares the strength of azimuthal scatter to radial abundance change for newly formed stars as a function of lookback time. The orange line shows the mean abundance change in radial abundance between $R = 0\,\mathrm{kpc}$ and R^*_{90}. The blue line shows the mean 360° azimuthal scatter of [Fe/H], which we averaged across $R = 2, 4, 6, 8$, and $10\,\mathrm{kpc}$, given the modest radial dependence. The dashed line shows the best fit to the evolution of the azimuthal scatter (see Section 3.4.9; Fig. 3.6 shows the fit to the radial change). The shaded regions show the $1-\sigma$ scatters across our 11 galaxies.

As Fig. 3.7 showed, the strength of the radial variations increased with time as the radial gradient steepened. Also, as Fig. 3.10 showed, the azimuthal scatter decreased with time at all radii. The point at which these cross identifies a transition epoch, at which newly formed stars transitioned from forming in a galaxy primarily dominated by azimuthal scatter to a disk primarily dominated by a radial gradient. This transition necessarily correlates with the gas disks transitioning to being rotationally dominated; Ma et al. (2017a) showed that strong radial gradients are only found in galaxies with a gas disk characterized by well-ordered rotation.

This is a critical transition period to characterize for chemical tagging, because it in effect identifies the maximum age of stars for which 1-D radial models for chemical tagging provide a good approximation. For all stars that formed prior to this transition age, their abundance was influenced more by their azimuthal location than their radius at birth.

Fig. 3.11 (right) shows a histogram of this transition time for each simulated galaxy, which spans ≈ 6.4 to $\approx 10.6\,\mathrm{Gyr}$ ago. The black line shows the cumulative distribution. The black arrow shows the mean transition time, $\approx 8\,\mathrm{Gyr}$ ago. This transition age for stars is slightly earlier than the transition age for gas (as presented in B21), which varied with

Correlation Metric	Total Gradient		Outer Gradient	
	Correlation	p-value	Correlation	p-value
$\sigma_{v_r}^{\rm star,young}/v_{\rm circ}$	0.727	0.011	0.836	0.001
$\sigma_{v_r}^{\rm star,young}$	0.655	0.011	0.764	0.006
Median stellar age	-0.745	0.008	-0.682	0.021
j_z/j_c	-0.409	0.212	-0.509	0.110
Bursty to smooth SFR time	-0.464	0.151	-0.427	0.19
Transition lookback time	-0.351	0.290	-0.469	0.145
$\Sigma^{\rm star,young}/\Sigma^{\rm gas}$	0.127	0.709	0.336	0.312
$\nabla\left(\Sigma^{\rm star,young}/\Sigma^{\rm gas}\right)$	0.027	0.937	-0.291	0.385
$R_{90}^{\rm star,all}$	0.223	0.509	-0.036	0.916
$R_{90}^{\rm star,young}$	-0.005	0.989	-0.196	0.564
$\nabla\Sigma^{\rm star,young}$	0.027	0.937	-0.127	0.709

TABLE 3.2. The Spearman rank correlation coefficient and corresponding p-value between the total stellar [Fe/H] gradient or the outer stellar [Fe/H] gradient in young stars (age $< 500\,$Myr) and different metrics of the galaxies' formation histories. The metrics we use, ranked by the average p-value of the total and outer gradient correlations, are: the ratio of the radial velocity dispersion of young stars to their average circular velocity ($\sigma_{v_r}^{\rm star,young}/v_{\rm circ}$), the radial velocity dispersion of young stars ($\sigma^{\rm star,young}$), the median age of stars in the galaxy, the circularity parameter of young stars j_z/j_c (Abadi et al., 2003), the transition time from bursty to smooth star formation in Yu et al. (2021), the transition time from azimuthal scatter domination to radial gradient domination in Section 3.4.7, the break radius of a two-component linear profile fit to the ratio of young stellar surface density to gas surface density shown in Fig. B.1, the change in the surface density ratio of young stars to gas over the same radial range divided by $R_{90} - 3\,$kpc ($\nabla\Sigma^{\rm star,young}/\Sigma^{\rm gas}$), the $R_{90}^{\rm star,all}$ and $R_{90}^{\rm star,young}$ in Section 3.4.1, and the change in the surface density of young stars from $R = 3\,$kpc to R_{90} divided by $R_{90} - 3\,$kpc ($\nabla\Sigma^{\rm star,young}$). We find the only correlations with significant p-values to be with: median age, $\sigma^{\rm star,young}$, and $\sigma_{v_r}^{\rm star,young}/v_{\rm circ}$.

radius from 7.4 Gyr ago at $R = 4\,$kpc to 6.9 Gyr ago at $R = 12\,$kpc (see Appendix B.2 for more discussion).

A potentially important caveat to applying this result to the MW is that, as Fig. 3.5 showed, our simulated gradients at $z = 0$ are likely shallower than the MW. If this discrepancy persisted across time, then the gradient of the MW was steeper than these simulations predict, which suggests a potentially earlier transition age, $\gtrsim 10.8\,$Gyr. However, we caution

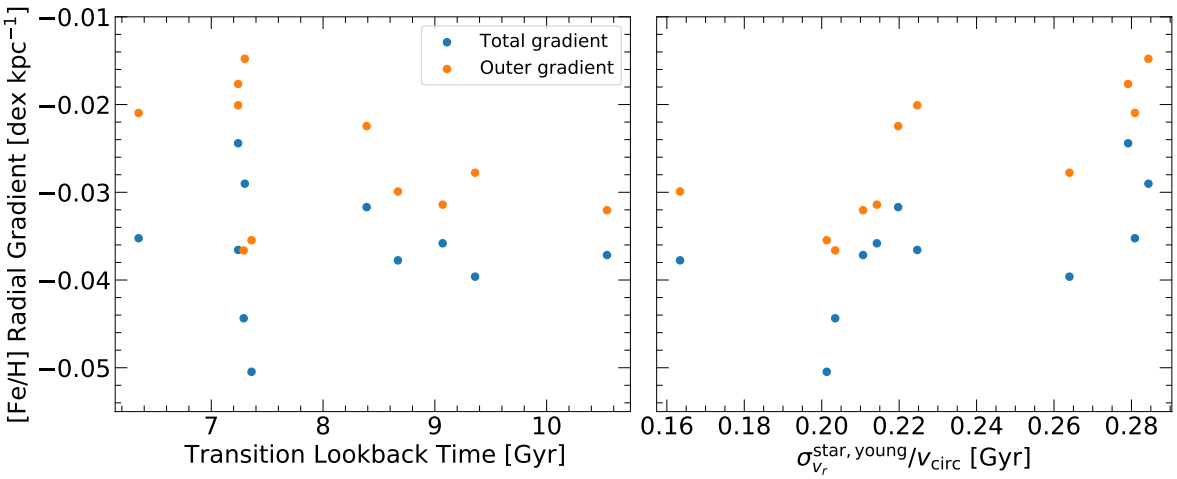


FIGURE 3.12. Correlation of the radial gradient in stars younger than 500 Myr against various galaxy properties. The blue points show the total gradient and the orange points show the outer gradient, as discussed in Section 3.4.3. **Left**: The radial gradient versus the transition time from Fig. 3.11 (right). While galaxies that transitioned at earlier lookback times have slightly steeper gradients at $z = 0$, on average, the correlations are not statistically significant (see Table 3.2). As Table 3.2 shows, the strength of the radial gradient at $z = 0$ has little to no significant correlation with any metric of disk 'settling' time, although it does show a reasonably strong correlation with overall stellar age. **Right**: The strongest correlation with the radial abundance gradients at $z = 0$ is with $\sigma_{v_r}^{\text{star,young}}/v_{\text{circ}}$. Galaxies that are more rotationally dominated have the stronger radial gradients in abundance, likely because of less radial mixing.

that the strength of the gradient at $z = 0$ does not necessarily correlate strongly with its behavior many Gyrs ago, as we show below.

3.4.8. What determines the present-day radial gradient? To understand better what aspect of formation history determines, or at least correlates with, the strength of the radial abundance gradient at $z = 0$, we calculate the Spearman rank correlation coefficient of the radial gradient (both total and outer) of the youngest stars in our galaxies with a variety of different metrics. Table 3.2 shows all metrics and associated correlations and p-values.

Fig. 3.12 shows example scatter plots for two correlation metrics. Blue points show the slope of the total gradient and the orange points show the slope of the outer gradient. The left panel shows the correlation of the [Fe/H] gradients with transition lookback time

(described in Section 3.4.7). The right panel shows the most significant correlation we find, the correlation between [Fe/H] gradients and the ratio of radial velocity dispersion to circular velocity for stars younger than 500 Myr ($\sigma_{v_r}^{\rm star,young}/v_{\rm circ}$).

We find no significant correlation between the radial gradients and the transition time from bursty to smooth star formation in these galaxies (as presented in Yu et al., 2021), nor with the transition times we present in Section 3.4.7. This indicates that the steepness of the radial abundance gradient does not depend on any metric of when the disk 'settled', such as the amount of time the galaxy has experienced smooth star formation or the time since the radial gradient became the dominant source of abundance inhomogeneity. We also find no significant correlation between the radial gradients and the size of the galaxy (measured using all or young stars), or with the strength of the gradient of the ratio of young-star to gas surface density (both measured from $R = 3\,$kpc to $R = R_{90}^{*}$ to exclude the bulge region). Although Fig. 3.4 shows general agreement between the *average* shape of $\Sigma^{\rm star,young}/\Sigma^{\rm gas}$ and the *average* abundance profile across our suite, there is significant host-to-host scatter.

We also find no correlation between the radial gradient and the circularity parameter (see Abadi et al., 2003) of young stars (age $\lesssim 500\,$Myr), defined as the ratio of the angular momentum of a star to the angular momentum with the same energy on a circular orbit.

We do find a statistically significant correlation of the radial gradient with the radial velocity dispersion of stars younger than 500 Myr ($\sigma_{v_r}^{\rm star,young}$) and with the ratio of radial velocity dispersion to circular velocity for young stars ($\sigma_{v_r}^{\rm star,young}/v_{\rm circ}$). These metrics of 'diskiness' at $z \sim 0$ imply that the dynamics of stars is an important factor in setting the radial gradient strength, even though the circularity parameter alone is not significant. In galaxies with significant radial velocity dispersion in young stars, there must be more radial mixing in gas (a newly formed star particles takes on the kinematic properties of its progenitor gas cell). This causes the gradients to be shallower. This implies that the degree

of radial mixing (via turbulence, spiral arms, and so on) at $z \sim 0$ is more important in determining the radial gradient of abundances than the formation history.

That said, we do find one significant correlation with a formation-history metric: a negative correlation with the median stellar age of the galaxy. This implies that galaxies whose stars formed earlier have steeper gradients at present day. This agrees with previous analysis of low mass galaxies in FIRE (for example Mercado et al., 2021). But interestingly this correlation does not extend to any metric of disk settling time. In other words, we find a correlation with when the stars formed, but not with when they formed in a settled disk. We defer a more detailed analysis of the relationship between disk settling time and star-formation history to future work.

3.4.9. Fits to functional forms. For stars at the time of their formation, we quantify the evolution of the galaxy-wide abundance across time (Fig. 3.1 middle), the overall radial gradient across time (Fig. 3.7 top), and the 360° azimuthal scatter across time (Fig. 3.11 left). We fit each galaxy independently, as well as the mean trends across all 11 galaxies. We tested fitting as a function of redshift, expansion scale factor, and stellar age. We found fitting as a function of stellar age to provide the best fits.

Table 3.3 shows the best fit to the average [Fe/H] and [Mg/H] of all stars at formation within the galaxy across time, that is, the fit to Fig. 3.1 (middle). We fit to the functional form:

$$[\mathrm{X/H}] = A - Be^{t_{\mathrm{lb}}/\tau} \tag{3.2}$$

We determine the best fit coefficients using the optimize.curve_fit function in SciPy. We tested second-order polynomial fits as well as simple exponential fits, but we find better agreement with the functional form in Eq. 3.2.

sim	[Fe/H]			[Mg/H]		
	A [dex]	B [dex]	τ [Gyr]	A [dex]	B [dex]	τ [Gyr]
m12m	3.1×10^{-1}	1.0×10^{-1}	4.0	5.0×10^{-1}	6.9×10^{-2}	3.7
Romulus	1.5×10^{-1}	5.0×10^{-2}	3.6	3.3×10^{-1}	2.8×10^{-2}	3.2
m12b	7.4×10^{-2}	1.1×10^{-2}	2.4	2.8×10^{-1}	5.6×10^{-3}	2.2
m12f	1.6×10^{-2}	1.5×10^{-2}	2.6	2.4×10^{-1}	9.3×10^{-3}	2.4
Thelma	3.8×10^{-1}	2.8×10^{-1}	6.0	5.6×10^{-1}	2.2×10^{-1}	5.5
Romeo	6.8×10^{-2}	1.4×10^{-2}	2.7	2.5×10^{-1}	5.2×10^{-3}	2.3
m12i	1.2×10^{-1}	5.8×10^{-2}	3.6	3.2×10^{-1}	4.0×10^{-2}	3.3
m12c	2.5×10^{-1}	1.5×10^{-1}	5.1	4.3×10^{-1}	1.1×10^{-1}	4.6
Remus	8.1×10^{-2}	4.8×10^{-2}	3.8	2.4×10^{-1}	2.3×10^{-2}	3.2
Juliet	9.9×10^{-3}	1.8×10^{-2}	2.9	1.9×10^{-1}	9.3×10^{-3}	2.5
Louise	5.2×10^{-2}	8.8×10^{-2}	4.4	2.0×10^{-1}	5.1×10^{-2}	3.8
Mean	1.1×10^{-1}	5.4×10^{-2}	3.6	3.0×10^{-1}	3.3×10^{-2}	3.2

TABLE 3.3. Fits for each simulation to stellar [Fe/H] and [Mg/H] (shown in Fig. 3.1) versus age, using newly formed, in-situ stars across the galaxy. We fit the profile of each host to [X/H] $= A - Be^{t_{\rm lb}/\tau}$, where $t_{\rm lb}$ is stellar age (lookback time). The bottom row shows the fit to the mean across these 11 galaxies.

Table 3.4 shows the best fit to the time evolution of the overall radial gradient, $\Delta[\mathrm{X/H}]_{R^*_{90}}/R^*_{90}$. We fit a second-order polynomial:

$$\Delta[\mathrm{X/H}]_{R^*_{90}}/R^*_{90} = At^2_{\rm lb} + Bt_{\rm lb} + C \tag{3.3}$$

We determine the best fit coefficients using the polyfit function from NumPy. We select a second-order fit, because the radial gradient evolution of some galaxies is too complicated to be captured by a linear fit, and a 3rd-order polynomial suffers from overfitting.

Table 3.5 shows the best fits to the evolution of the 360° azimuthal scatter of [Fe/H]. We average the azimuthal scatter over 5 radii (2, 4, 6, 8 , and 10 kpc). We fit a 2nd order polynomial using the polyfit function in NumPy, as with the radial gradient evolution:

$$\sigma^{360^\circ}_{\rm [X/H]} = At^2_{\rm lb} + Bt_{\rm lb} + C \tag{3.4}$$

We tested an exponential fit too, but we find better agreement for a second-order polynomial.

sim	[Fe/H]			[Mg/H]		
	A $\left[\frac{\text{dex}}{\text{kpc Gyr}^2}\right]$	B $\left[\frac{\text{dex}}{\text{kpc Gyr}}\right]$	C $\left[\frac{\text{dex}}{\text{kpc}}\right]$	A $\left[\frac{\text{dex}}{\text{kpc Gyr}^2}\right]$	B $\left[\frac{\text{dex}}{\text{kpc Gyr}}\right]$	C $\left[\frac{\text{dex}}{\text{kpc}}\right]$
m12m	-4.5×10^{-5}	3.1×10^{-3}	-3.5×10^{-2}	-4.5×10^{-5}	1.9×10^{-3}	-3.0×10^{-2}
Romulus	-1.2×10^{-4}	2.6×10^{-3}	-3.4×10^{-2}	-1.4×10^{-7}	1.1×10^{-3}	-2.6×10^{-2}
m12b	1.3×10^{-3}	-1.2×10^{-2}	-3.5×10^{-2}	1.2×10^{-3}	-1.2×10^{-2}	-2.6×10^{-2}
m12f	4.1×10^{-4}	-2.3×10^{-3}	-3.0×10^{-2}	3.6×10^{-4}	-2.1×10^{-3}	-2.4×10^{-2}
Thelma	3.8×10^{-4}	-8.7×10^{-4}	-2.4×10^{-2}	6.1×10^{-4}	-3.3×10^{-3}	-1.8×10^{-2}
Romeo	3.3×10^{-4}	-1.1×10^{-3}	-3.8×10^{-2}	4.0×10^{-4}	-2.5×10^{-3}	-2.7×10^{-2}
m12i	3.6×10^{-4}	9.7×10^{-6}	-3.6×10^{-2}	3.7×10^{-4}	-7.7×10^{-4}	-2.8×10^{-2}
m12c	1.4×10^{-4}	2.1×10^{-3}	-3.7×10^{-2}	2.2×10^{-4}	7.8×10^{-4}	-3.0×10^{-2}
Remus	3.2×10^{-4}	-1.3×10^{-3}	-3.6×10^{-2}	3.1×10^{-4}	-1.7×10^{-3}	-2.8×10^{-2}
Juliet	-1.9×10^{-4}	6.3×10^{-3}	-5.5×10^{-2}	-1.1×10^{-4}	4.6×10^{-3}	-4.3×10^{-2}
Louise	3.8×10^{-4}	2.9×10^{-4}	-4.3×10^{-2}	4.9×10^{-4}	-9.6×10^{-4}	-3.6×10^{-2}
Mean	3.0×10^{-4}	-3.0×10^{-4}	-3.7×10^{-2}	3.6×10^{-4}	-1.3×10^{-3}	-2.9×10^{-2}

TABLE 3.4. Fits to the radial gradient ($\Delta[\text{X/H}]_{R^*_{90}}/R^*_{90}$) of stars at formation versus lookback time for each simulation, for both [Fe/H] and [Mg/H]. We fit to a second-order polynomial: $\Delta[\text{X/H}]_{R^*_{90}}/R^*_{90} = At^2_{\text{lb}} + Bt_{\text{lb}} + C$. The bottom shows the fit to the mean trend across all 11 galaxies.

sim	[Fe/H]			[Mg/H]		
	A $\left[\frac{\text{dex}}{\text{Gyr}^2}\right]$	B $\left[\frac{\text{dex}}{\text{Gyr}}\right]$	C [dex]	A $\left[\frac{\text{dex}}{\text{Gyr}^2}\right]$	B $\left[\frac{\text{dex}}{\text{Gyr}}\right]$	C [dex]
m12m	1.3×10^{-3}	-1.1×10^{-3}	3.3×10^{-2}	1.6×10^{-3}	-5.3×10^{-3}	3.2×10^{-2}
Romulus	9.8×10^{-4}	-6.0×10^{-3}	4.4×10^{-2}	1.0×10^{-3}	-5.3×10^{-3}	3.5×10^{-2}
m12b	6.7×10^{-4}	-3.0×10^{-3}	4.6×10^{-2}	7.3×10^{-4}	-2.5×10^{-3}	3.8×10^{-2}
m12f	1.4×10^{-3}	-3.7×10^{-3}	7.2×10^{-2}	1.4×10^{-3}	-2.7×10^{-3}	6.0×10^{-2}
Thelma	1.2×10^{-3}	-2.5×10^{-4}	3.3×10^{-2}	1.2×10^{-3}	-3.1×10^{-4}	2.8×10^{-2}
Romeo	1.7×10^{-3}	-8.5×10^{-3}	6.6×10^{-2}	1.7×10^{-3}	-7.5×10^{-3}	5.1×10^{-2}
m12i	1.4×10^{-3}	-1.0×10^{-2}	4.8×10^{-2}	1.4×10^{-3}	-8.9×10^{-3}	4.1×10^{-2}
m12c	7.6×10^{-4}	-3.6×10^{-3}	4.4×10^{-2}	9.0×10^{-4}	-4.2×10^{-3}	4.0×10^{-2}
Remus	5.3×10^{-4}	-1.6×10^{-3}	4.0×10^{-2}	5.7×10^{-4}	-9.3×10^{-4}	3.2×10^{-2}
Juliet	1.6×10^{-3}	-1.2×10^{-2}	9.4×10^{-2}	1.4×10^{-3}	-8.3×10^{-3}	7.7×10^{-2}
Louise	1.5×10^{-3}	-5.2×10^{-3}	4.9×10^{-2}	1.7×10^{-3}	-5.7×10^{-3}	4.2×10^{-2}
Mean	1.2×10^{-3}	-4.6×10^{-3}	5.2×10^{-2}	1.2×10^{-3}	-4.1×10^{-3}	4.3×10^{-2}

TABLE 3.5. Fits to the total (360°) azimuthal scatter of stars at formation, averaged across all radii, versus lookback time for each simulation. We fit a second-order polynomial $\sigma^{360^\circ}_{\text{[X/H]}} = At^2_{\text{lb}} + Bt_{\text{lb}} + C$. The bottom shows the fit to the mean trend across all 11 galaxies.

3.4.10. Comparison to previous results for all gas. B21 examined the evolution of elemental abundance variations for all gas (not just star-forming gas) in the same MW-mass FIRE-2 simulations. Here, we expand on those results by specifically examining newly formed stars. We generally expect the trends for newly formed stars to match those of the gas, and while we find overall qualitatively similar results, we briefly summarize key quantitative differences and similarities.

For radial gradients, we find generally steeper radial gradients in newly formed stars relative to all gas. For vertical gradients, both all gas and newly formed stars exhibit little to no variation with height. For azimuthal scatter, newly formed stars are weaker than that of all gas. This indicates a higher degree of homogeneity for star-forming gas. Additionally, Fig. 3.10 shows that the azimuthal scatter of abundances for newly formed stars is independent of radius. This contrasts the results of B21, who found that the azimuthal scatter of all gas increases with increasing radius. Appendix B.2 directly compares the

evolution of azimuthal scatter and radial change in abundance for newly formed stars versus all gas, following the methods of B21.

Following the analysis of B21 (their Section 3.5), Appendix. B.3 shows the radial scale at which azimuthal abundance variations are subdominant to the radial variations in abundance ($\Delta R_{\rm equality}$) and the radial scale at which radial variations are measurable homogeneous ($\Delta R_{\rm homogeneous}$. $\Delta R_{\rm equality}$ and $\Delta R_{\rm homogeneous}$ for newly formed stars are smaller than for all gas, which reflects the steeper radial gradients and smaller azimuthal scatter in newly formed stars.

3.5. Summary and Discussion

3.5.1. Summary. We used a suite of 11 MW/M31-mass cosmological zoom-in simulations, run with FIRE-2 physics, to explore the 3-D spatial variations of elemental abundances of stars *at birth* (within $< 50\,$Myr of their formation) across these galaxies' formation histories. (In future work we will examine similar trends for stellar populations at $z = 0$ as a function of their age.) We measured properties of newly formed stars as a function of lookback time going back $\approx 12\,$Gyr, in part to test and guide approaches to chemical tagging. We also fit functional forms to these trends, to use in models of elemental evolution. Our main results are:

- *Galaxy stellar abundances*: enrich relatively quickly: 5 of our 11 galaxies reached [Fe/H] ≈ -0.5 at lookback times of $9 - 10\,$Gyr ago. LG-like galaxies enriched in metals faster than isolated galaxies, following their more rapid stellar mass assembly (see Fig. 3.1 as well as Santistevan et al. 2020).
- *Galaxy stellar size*: R^*_{90} for both young stars and all stars increased over time. R^*_{90} for all stars was comparable to that of young stars $\gtrsim 7.5\,$Gyr ago. However, after this, R^*_{90} of young stars is systematically larger than that of all stars, reflecting inside-out radial growth. For $t_{\rm lb} \lesssim 8\,$Gyr, using either all stars or young stars, R^*_{90} is larger for galaxies in LG-like environments than those that are isolated ($16.2\,$kpc

versus 13.1 kpc for young stars and 11.1 kpc versus 10.2 kpc for all stars at $z = 0$; see also Garrison-Kimmel et al. 2018).

- *Galaxy-wide scatter in abundances*: reached a minimum of ≈ 0.09 dex for in-situ stars forming ≈ 7 Gyr ago. This reflects a competition between a reduction of the scatter as azimuthal variations decreased over time and an increase in the galaxy-wide scatter as the radial gradient became stronger.
- *Vertical gradients*: at formation are negligible in nearly all regimes. The change in abundance is on average less than 0.02 dex over 1 kpc for both [Fe/H] and [Mg/H]. Thus, vertical abundance variations provide minimal discriminating power for chemical tagging. The one exception that we find is for stars in the inner bulge region, $R \lesssim 4$ kpc, at $z \approx 0$.
- *Radial gradients*: of newly formed stars were flat (magnitude $\lesssim 0.01\,\mathrm{dex\,kpc^{-1}}$) at lookback times $\gtrsim 9.5$ Gyr but became progressively steeper with time, reaching $-0.037\,\mathrm{dex\,kpc^{-1}}$ for [Fe/H] ($-0.030\,\mathrm{dex\,kpc^{-1}}$ for [Mg/H]) at $z = 0$. [Fe/H] gradients for newly formed stars at $z = 0$ are shallower than in the MW measured over similar radial ranges. However, B21 showed that gas-phase radial abundance gradients in our simulations are as steep or steeper than those observed in external MW-mass galaxies. Our galaxies are well fit by two-component radial gradients that are steeper in the inner galaxy, which reflects a steeper stellar surface density profile in the inner galaxy, though not necessarily with the bulge region.
- *Azimuthal scatter*: of young stars systematically decreases over time from $\lesssim 0.18$ dex 11.6 Gyr ago to $\lesssim 0.043$ dex today for [Fe/H] and [Mg/H]. Azimuthal scatter shows minimal dependence on azimuthal bin size, so small-scale variations dominate over larger-scale variations. Even at scales $\lesssim 1$ kpc we measure [Fe/H] azimuthal scatter in the solar cylinder of ≈ 0.043 dex at $z = 0$ and ≈ 0.16 dex for stars that formed at $t_{\rm lb} = 11.6$ Gyr. Importantly, our analysis does *not* center on individual star-forming

regions, but rather, random patches in the galaxy, so this is not a statement about the internal homogeneity of individual star-forming regions and star clusters.

- *Azimuthal versus radial variations*: Similar to our analysis of gas in B21, our simulated galaxies transitioned from being dominated by azimuthal scatter to being dominated by radial variations at lookback times of $\approx 8\,\mathrm{Gyr}$ ago. Thus, azimuthal variations were the primary source of galaxy-wide scatter in abundance at early times, and they are of secondary importance (though not negligible) for stars that formed $\lesssim 8\,\mathrm{Gyr}$ ago.
- *Correlations with present-day radial gradient*: We tested the correlation between the strength of the radial gradient at $z = 0$ and a variety of metrics of formation history. The most statistically significant correlation is the ratio of the radial velocity dispersion of young stars to their circular velocity. So, the degree of radial mixing in galaxies is likely set by the strength of the ratio of radial velocity dispersion to circular velocity. Additionally, we find a lack of correlation with the transition lookback time (see Section 3.4.7) which implies disk settling time is not responsible for setting present-day abundance gradients.
- *Fit to functional forms*: We fit the evolution of overall normalization, the radial gradient, and the azimuthal scatter of the abundances of stars at formation versus lookback time (stellar age) in Section 3.4.9.

3.5.2. Limitations and caveats. The simulations analyzed in this work implement the FIRE-2 physics, discussed in Section 3.3.1. Hopkins et al. (2018) presents the physics in detail and a variety of tests showing their robustness robustness. However there are still limitations inherent to our analysis and the physics implemented in FIRE-2.

We analyze only 11 galaxies, so our results are limited by our sample size. Furthermore, because we chose these galaxies to be near the mass of the MW, they necessarily encompass a narrow range of stellar and halo masses (see Table 3.1).

Also, as Sanderson et al. (2020); McCluskey et al. (2023) show, the velocity dispersion of stars in these simulations is dynamically hotter than observed in the MW (though they are more similar to M31). This could play a role in the shallower radial gradients in our simulations (see Section 3.4.3), given the strong correlation between stellar velocity dispersion and strength of the radial gradient (see Section 3.4.8).

In addition, there are limitations in our current physics implementations. The simulations do not include a self-consist treatment of cosmic rays or magnetohydrodynamics and anisotropic thermal conduction and viscosity in gas (see Hopkins et al., 2018). FIRE-2 treats all core-collapse supernovae as having identical IMF-averaged yields, but different mass progenitors will have different yields (see Muley et al., 2021). Future FIRE-3 simulations will appropriately mass sample rates and yields of different mass core-collapse supernovae (Hopkins et al., 2023). Additionally, Gandhi et al. (2022) showed that the default implementation of type Ia supernovae rates in FIRE-2 may be underestimated, leading to an underproduction of [Fe/H].

Finally, these FIRE-2 simulations do not include any treatment of AGN from supermassive black holes, which may bias the dynamics and star-formation rates, particularly in the inner few kpc. However, recent implementations in FIRE (Wellons et al., 2023) will allow us to explore their effects in future work.

3.5.3. Discussion. Our analysis extends the work in B21, in which we examined the homogeneity of gas, as an initial guide for the homogeneity of newly formed stars. The primary goal of this paper is to quantify the elemental abundance homogeneity of newly formed stars as a function of lookback time and to provide functional forms to its evolution such that chemical-tagging models have better descriptors of the initial degree of homogeneity with which stars form in a cosmological context. This will allow for more accurate galactic elemental evolution models and provide more realistic expectations for chemical tagging models.

We examined [Fe/H], [Mg/H], and [Mg/Fe]. We look at Fe primarily as a representative element of type Ia supernovae and Mg as a representative α element, that is, primarily sourced via core collapse supernovae. Because our analysis is limited to only 2 elements, we do not asses if analysis of more abundances will provide more discriminating power for chemical tagging. However, the work of Ting & Weinberg (2022) suggests including at least $7-8$ elements when doing galactic archaeology, Casamiquela et al. (2021) concluded that the larger the abundance space the better for chemical tagging, and Ratcliffe et al. (2022) found that stellar clusters are better identified when using 15 abundances rather than 2. Our FIRE-2 simulations track 9 metals. Given the expected correlation between elements primarily sourced via the same enrichment channels, we defer a more detailed analysis of all 9 elements to future work using the FIRE-3 simulations (Hopkins et al., 2023), which implement a tracer-element approach for varying stellar yields in post-processing (Wetzel et al., 2023).

As we showed in Section 3.4.1, the average [Fe/H] of stars in our galaxies as well as the average stellar mass of our galaxies are similar to those of the MW. Additionally, we measured the scatter in [Fe/H] as a function of stellar age and found it decreases with decreasing stellar age for stars older than $\approx 7\,\mathrm{Gyr}$ and then increases with decreasing stellar age. This is discrepant with previous analysis of stellar metallicity distributions in the MW that find the scatter in [Fe/H] continually decreases with decreasing stellar age (Casagrande et al., 2011; Miglio et al., 2021). However, we do not match the selection function of these surveys, nor do we divide our stars into similar age bins.

However, an important caveat to this work is that our simulations are not designed to recreate the history of the MW; they instead provide a cosmologically representative range of histories of galaxies that are similar to the MW at $z = 0$. Boardman et al. (2020) suggest that, because the MW has a particularly small disc scale length relative to similar mass galaxies, it is important to factor in disc scale length when selecting MW analogs. However,

in B21 we tested scaling the gas-phase radial abundance gradients in our MW-mass galaxies and found the most self-similarity when measuring gradients in physical units.

We explore the effect of varying the age range of stars used to measure the radial gradients in the simulations, because the gradients are systematically shallower than observations of the MW, although our gradients are steeper than those measured in nearby MW-mass galaxies. This in principle could account for uncertainties in the ages of observed star clusters. However, including older populations of stars leads to increasingly shallower gradients, resulting in greater discrepancies with observations. Some observations (for example Carrera & Pancino, 2011; Cunha et al., 2016; Netopil et al., 2016; Donor et al., 2020; Santos-Peral et al., 2021) find that the MW's galactic radial gradient in abundance as determined by older star clusters is steeper than the gradient as determined by younger star clusters, but our results imply the opposite. Our results agree with [O/H] measurements from planetary nebulae in M81 (for example Stanghellini et al., 2010, 2014) and in the MW (Stanghellini & Haywood, 2018). However, we do not measure the radial positions of stars at $z = 0$ as a function of age, although we plan to pursue this in future work. For our comparison we examine the effect of widening the age cutoff we use to identify 'young' stars at $z = 0$ from $< 0.5\,\mathrm{Gyr}$ to $< 10\,\mathrm{Gyr}$. The shallower radial gradients when including older stars likely results from radial mixing (for example Schönrich & Binney, 2009; Loebman et al., 2011; Quillen et al., 2018).

The steepening of the radial gradients in our simulations disagrees with some observations which indicate either a flattening of radial abundance gradients (for example Frinchaboy et al., 2013; Netopil et al., 2016; Spina et al., 2017) or a steepening and then flattening (for example Xiang et al., 2015, 2017; Anders et al., 2017) but agrees with others (for example Stanghellini & Haywood, 2010, 2018). However, present-day measurements of stellar positions do not necessarily represent the formation locations of stars, so these results are not directly comparable to our work. A better point of comparison is either spatially resolved

observations of stars in MW progenitor analogues, or spatially resolved gas-phase abundance observations (as discussed in B21), assuming the stars form with abundances representative of the gas. High-redshift observations of gas-phase abundance gradients generally show flat radial profiles (for example Wuyts et al., 2016; Patrício et al., 2019; Curti et al., 2020), but some show strong negative radial gradients (for example Carton et al., 2018; Wang et al., 2020). At high redshifts, galaxies do not sustain gas -phase abundance gradients (B21). Once disk settling occurs, gas no longer mixes radially as efficiencly in the disk, so it can take on different abundances at different radii, imprinted on newly formed stars.

In addition to our fiducial measure of the total radial gradient, we fit a two-component piecewise linear function to the abundance profile. We generally find that the inner region of the disk is steeper than the outer region, in contrast to some observations (for example Hayden et al., 2014; Maciel & Andrievsky, 2019), but in agreement with other observations (for example Netopil et al., 2016; Reddy et al., 2020). Exact comparisons with observations are difficult, because observations typically measure the radial gradient for all stars, rather than measuring for mono-age stellar populations. Also, uncertainties in the inferred ages and locations of stars can influence the interpreted gradients.

One of the key results of this paper is the quantification of the evolution of the azimuthal scatter in abundance of newly formed stars. The azimuthal scatter does not directly track the scatter in the gas in B21. The scatter is systematically smaller for stars than for all gas. This discrepancy is larger at larger lookback times and larger radii (see Fig. B.2) which results from stars at large lookback times and large radii preferentially forming from gas occupying slightly higher metallicity. This tests a common assumption in chemical-tagging models, that stars have abundances that primarily depend just on their birth radii and not on azimuthal position (for example Frankel et al., 2018, 2020).

Azimuthal scatter of stellar abundances is important to quantify for chemical tagging, but has not yet been well characterized by observations. Observations of cepheids in the

MW by Luck et al. (2011) indicate no significant azimuthal dependence, which is consistent with the small azimuthal scatter we measure for the youngest stars in Fig. 3.9. However, our results also indicate much smaller azimuthal scatter than that observed in the MW by Kovtyukh et al. (2022). Measuring cepheids at $R = 7 - 9\,\mathrm{kpc}$ Kovtyukh et al. (2022) find [Fe/H] varies by up to 0.2 dex, much larger than our measured 0.05 dex scatter.

Additionally, in this paper we quantify a transition lookback time, in the same way as B21. This transition lookback time sets the timescale over which elemental evolution models reasonably can assume azimuthal homogeneity (for example Minchev et al., 2018; Mollá et al., 2019a; Frankel et al., 2020). Notably, the lookback time for newly formed stars is $\sim 8\,\mathrm{Gyr}$, $\approx 1\,\mathrm{Gyr}$ earlier than the lookback time derived for the gas in B21. This is primarily because the azimuthal scatter in stars at large lookback times were smaller than that of all gas. This may reflect the disk-wide scatter in all gas being slightly smaller than the disk-wide scatter in star-forming gas (B21).

Complementary to our analysis is that of Yu et al. (2021), who measured the transition epoch from 'bursty' to 'steady' star formation and disk settling in the same simulations. Similar to the analysis presented in B21, we find that our transition times are consistently earlier than the transition times in Yu et al. (2021), by $\sim 3.2\,\mathrm{Gyr}$ on average. However, our transition lookback times are moderately correlated with those in Yu et al. (2021) (Pearson correlation coefficient $r \approx 0.63$). Thus, the onset of a strong radial gradient in abundance precedes but correlates with the onset of 'steady' star formation in these simulations.

Perhaps most important to this analysis is our characterization of the evolution of elemental abundances, radial gradients, and azimuthal scatter across our suite of MW-mass simulations (see Section 3.4.9). We provide simple functional forms which encapsulate the evolution of abundance distributions of MW-mass galaxies across time. These functional

forms and fits are *crucial* to the future of elemental-evolution modeling and accurate chemical tagging, because weak chemical tagging is infeasible without an accurate picture of the birth conditions of stars.

Acknowledgements

We thank the anonymous reviewer for the detailed comments which we feel have significantly improved the quality and clarity of the article.

We performed this work using the GizmoAnalysis package (Wetzel & Garrison-Kimmel, 2020), the Astropy package (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018), as well as libraries from Numpy (Harris et al., 2020), SciPy (Virtanen et al., 2020), and Matplotlib (Hunter, 2007).

MB and AW received support from: the NSF via CAREER award AST-2045928 and grant AST-2107772; NASA ATP grants 80NSSC18K1097 and 80NSSC20K0513; HST grants AR-15809, GO-15902, GO-16273 from STScI; a Scialog Award from the Heising-Simons Foundation; and a Hellman Fellowship. We performed this work in part at the Aspen Center for Physics, supported by NSF grant PHY-1607611. We ran simulations using: XSEDE, supported by NSF grant ACI-1548562; Blue Waters, supported by the NSF; Frontera allocations AST21010 and AST20016, supported by the NSF and TACC; Pleiades, via the NASA HEC program through the NAS Division at Ames Research Center.

Data Availability

The data in these Figures/Chapter3 are available at `https://mbellardini.github.io/`. The FIRE-2 simulations are publicly available (Wetzel et al., 2023) at `http://flathub.flatironinstitute.org/fire`. Additional FIRE simulation data is available at `https://fire.northwestern.edu/data`. A public version of the GIZMO code is available at `http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html`.

CHAPTER 4

Measuring stellar motion: Radial and angular momentum redistribution of stellar orbits in FIRE simulations of Milky Way-mass galaxies with varying definitions of orbital radius

4.1. Abstract

A key question in galactic dynamics, which is central to most approaches to galactic archeology, is: how much do the orbits of stars change (radially redistribute) after their birth? We use cosmological zoom-in simulations of 11 Milky Way-mass galaxies from the Feedback In Realistic Environment (FIRE-2) suite to quantify the change in the orbital angular momentum and radius of stars, and its dependence on age, present radius, and formation radius. We explore and compare 5 different ways of measuring orbital radius; we find generally consistent trends, but only when measuring orbital radius today and radial redistribution self-consistently. Stars younger than $\approx 6\,\mathrm{Gyr}$ generally lost angular momentum, while stars older than $\approx 6\,\mathrm{Gyr}$ generally gained angular momentum; this transition occurred more recently at smaller radii. The median inward/outward change in radius, ΔR, of a stellar population today is generally mild ($\lesssim 1\,\mathrm{kpc}$), though it is more significant and positive for the oldest stars. The scatter of this radial change, $\sigma(\Delta R)$, a standard metric of radial redistribution, increases with age up to $\approx 2-3\,\mathrm{Gyr}$ but saturates at $\approx 2-3\,\mathrm{kpc}$ for stars older than this, in contrast with common assumptions for radial redistribution. $\sigma(\Delta R)$ increases only mildly with radius today. Finally, we examine correlations of radial redistribution with galaxy properties, finding no correlation between $\sigma(\Delta R)$ and the time of disk onset, but significant correlation with the stellar mass and radius of the galaxy today.

4.2. Introduction

Throughout the lifetime of a star, the radius of its orbit can change significantly. Depending on the driving physical process this is sometimes referred to as radial migration (e.g. Lynden-Bell & Kalnajs, 1972; Schönrich & Binney, 2009; Loebman et al., 2011), churning/cold torquing, or blurring/kinematic heating (e.g Sellwood & Binney, 2002; Daniel et al., 2019). 'Churning' refers to systematic changes to the orbital angular momentum of stars which changes the guiding center radius of the star while leaving the 'random' component of its orbital energy unchanged (e.g. Grand et al., 2012). 'Blurring' describes oscillatory motion of stars about their guiding center radius without changes to their specific angular momentum.

There are a variety of physical processes responsible for altering the radial position of stars such as scattering off of non-axisymmetric features like galactic bars (e.g. Halle et al., 2018; Khoperskov et al., 2020) or spiral arms (e.g. Sellwood & Binney, 2002; Roškar et al., 2008b; Loebman et al., 2016), interactions with merging satellites (e.g. Quillen et al., 2009; Carr et al., 2022), or scattering off of giant molecular clouds (e.g. Schönrich & Binney, 2009). Understanding the radial redistribution of stars has increasingly gained attention in theoretical, (e.g. Sellwood & Binney, 2002; Daniel & Wyse, 2015) and numerical work (e.g. Brunetti et al., 2011; Minchev et al., 2011; Vera-Ciro et al., 2014; Halle et al., 2015, 2018; Vincenzo & Kobayashi, 2020) as well as observational studies (e.g. Feltzing et al., 2020; Vickers et al., 2021; Zhang et al., 2021; Lian et al., 2022; Carr et al., 2022; Lu et al., 2022).

This mixing of stellar populations across radii leaves complex imprints in elemental abundance space such as dispersion in the age-metallicity relation (e.g. Haywood, 2008; Casagrande et al., 2011; Carrillo et al., 2023), skewness of the metallicity distribution as a function of galactocentric radius (e.g. Hayden et al., 2015; Loebman et al., 2016; Martinez-Medina et al., 2016), and supersolar metallicity stars in the solar neighborhood (Kordopatis et al., 2015; Anders et al., 2017; Feuillet et al., 2018). Observational studies confirming the

role of radial redistribution are limited given the inability to directly measure the birth location of a star. Instead, observational techniques must rely upon chemical tagging (Freeman & Bland-Hawthorn, 2002) to characterize stellar radial redistribution.

Chemical tagging is a tool of galactic archaeology, the aim of which is to infer the birth location of stars using stellar abundances in conjunction with models of galactic chemical evolution. Stellar abundances are, to first order, invariant properties, directly linking stellar birth location with their location at $z = 0$. Given sufficiently precise measurements of stellar abundances and sufficiently accurate models of galactic chemical evolution, one can infer the birth location of a star given its age and elemental composition.

There have been a multitude of observational surveys with the goal of mapping out stellar elemental abundances for chemical tagging; such as GALactic Archaeology with Hermes (De Silva et al., 2015; Buder et al., 2018), the Large Area Multi-Object Fiber Spectroscopic Telescope (Cui et al., 2012), and the Apache Point Galactic Evolution Experiment (Majewski et al., 2017; Ahumada et al., 2020; Jönsson et al., 2020). Theoretical works have forged ahead, alongside these observational studies, to better characterize galactic chemical evolution (e.g. Minchev et al., 2018; Spitoni et al., 2019; Mollá et al., 2019a; Hemler et al., 2021; Frankel et al., 2020; Bellardini et al., 2021, 2022).

Typically, models of galactic chemical evolution presume stars form on circular orbit with with an abundance set by their birth radius (e.g. Minchev et al., 2018; Frankel et al., 2018, 2019, 2020; Sharma et al., 2021). This is well motivated by observational evidence of strong negative radial abundance gradients in the Milky Way (MW) (e.g Davies et al., 2009; Anders et al., 2014; Balser et al., 2015; Wenger et al., 2019) and in external galaxies (e.g. Pilyugin et al., 2014; Sánchez-Menguiano et al., 2016; Belfiore et al., 2017; Poetrodjojo et al., 2018). However, recent observational efforts have recorded azimuthal variations in abundances in the MW (e.g. Balser et al., 2011; Wenger et al., 2019) and in external galaxies (e.g Ho et al., 2018; Kreckel et al., 2019, 2020; Sánchez-Menguiano et al., 2020; Metha et al., 2022; Li et al.,

2023). Azimuthal abundance variations add another layer of complexity to chemical tagging as they break the assumption of stellar abundances being uniquely determined by their birth radius alone.

Recent theoretical works (e.g. Spitoni et al., 2019; Mollá et al., 2019b; Bellardini et al., 2021, 2022) have explored the time evolution of azimuthal abundance variations. While accurate characterization of their time evolution is important for modeling the initial conditions for chemical tagging it is likely insufficient to break degeneracies associated with stellar birth locations. Characterizing the typical scale of radial redistribution of stars as a function of present-day observables can provide a powerful prior for chemical tagging models.

In this paper we use simulations of MW/M31-mass galaxies from the FIRE-2 cosmological simulations to explore radial redistribution of stars. In Section. 4.3 we describe the simulations used in this analysis and our definitions of stellar radius. In Section. 4.4 we characterize changes in stellar angular momentum as a function of stellar age and changes in stellar radius as a function of stellar age and location. In Section. 4.5 we summarize the main results from this paper and discuss their implications.

4.3. Methods

4.3.1. Simulations. For this analysis, we use cosmological zoom-in simulations from the Feedback In Realistic Environments (FIRE) project[1] (Hopkins et al., 2018). Our simulation suite includes 5 isolated MW/M31-mass galaxies from the *Latte* suite (Wetzel et al., 2016) and 6 Local Group (LG)-like MW+M31 pairs from the 'ELVIS on FIRE' suite (Garrison-Kimmel et al., 2019a,b), which have halo masses $M_{200\mathrm{m}} = 1 - 2 \times 10^{12} \mathrm{M}_{\odot}$[2]. Following Bellardini et al. (2021) and Bellardini et al. (2022) we exclude 3 galaxies from the *Latte* suite: m12r and m12z due to their low stellar masses and m12w due to its compact gaseous disk.

[1]FIRE project web site: http://fire.northwestern.edu

[2]$M_{200\mathrm{m}}$ is defined as the total mass within the radius within which the mean density is 200 times the mean matter density of the universe

The initial baryon particle mass in the *Latte* simulations is $7070 M_\odot$ ($\approx 5000 M_\odot$ at $z = 0$ due to stellar mass loss), and the dark-matter mass resolution is $3.5 \times 10^5 M_\odot$. This is $\sim 2\times$ better in the ELVIS simulations: the Thelma & Louise simulation has initial baryon particle masses of $4000 M_\odot$ and the Romeo & Juliet and Romulus & Remus simulations have initial baryon particle masses of $3500 M_\odot$. The gravitational force softening of star and dark-matter particles is fixed with a Plummer equivalent of $\epsilon_{\rm star} = 4\,{\rm pc}$ and $\epsilon_{\rm dm} = 40\,{\rm pc}$. Gas cells have fully adaptive softening, which matches the hydrodynamic kernel smoothing, reaching a minimum softening length of 1 pc.

All simulations were run with the Meshless Finite Mass hydrodynamics method of GIZMO (Hopkins, 2015) and with star formation, stellar feedback, and fluid dynamics implementation of FIRE-2 (Hopkins et al., 2018). GIZMO enables adaptive hydrodynamic smoothing of gas elements based on their density while simultaneously conserving mass, energy, and momentum of particles to machine accuracy.

The FIRE-2 model incorporates physically motivated models of star formation. Star formation occurs only in gas that is self-gravitating, Jeans-unstable, cold ($T < 10^4$K), and molecular (following Krumholz & Gnedin, 2011). Individual star particles inherit the mass of their progenitor gas cell and represent single stellar populations, assuming a Kroupa (2001) stellar initial mass function. They evolve along standard stellar population models.

The FIRE-2 simulations also model time-resolved stellar feedback processes such as continuous mass loss from stellar winds, core-collapse and Ia supernovae, radiation pressure, photoionization, and photo-electric heating. We model stellar winds and their yields following a combination of models (Van den Hoek & Groenewegen, 1997; Marigo, 2001; Izzard et al., 2004) synthesized in Wiersma et al. (2009). Core-collapse and Ia supernovae rates come from STARBURST99 (Leitherer et al., 1999) and Mannucci et al. (2006) respectively. We use Nomoto et al. (2006) and Iwamoto et al. (1999) for the nucleosynthetic yields of core-collapse and Ia supernovae, respectively.

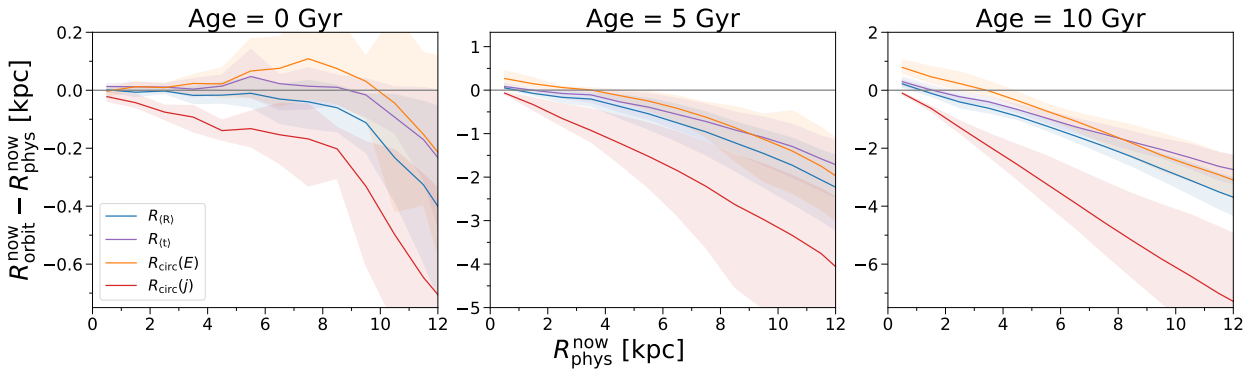


FIGURE 4.1. **Comparing 4 metrics of the orbital radius of stars at $z = 0$ against their instantaneous physical radius.** The lines show the mean across 11 galaxies (9 galaxies for $R_{\rm circ}(E)$ and $R_{\rm circ}(j)$) and the shaded regions show the 68th percentile scatter. The panels show different age stellar populations, corresponding (from left to right) to the late-disk, early-disk, and pre-disk eras. The different orbital radii agree best at smaller $R_{\rm phys}$ and for younger stars, while they are consistently smaller than $R_{\rm phys}$ at large $R_{\rm phys}$ and at older ages. In general, $\langle R_{\rm R} \rangle$, $\langle R_{\rm t} \rangle$, and $R_{\rm circ}(E)$ yield similar results, while $R_{\rm circ}(j)$ is generally the smallest.

Gas cells are subject to metallicity-dependent radiative heating and cooling processes (across a temperature range of $10 - 10^{10}$K) including free-free, photoionization and recombination, Compton, photo-electric and dust collisional, cosmic ray, molecular, metal-line, and fine structure processes, accounting for 11 elements (H, He, C, N, O, Ne, Mg, Si, S, Ca, Fe). All of the simulations include cosmic ultraviolet background following Faucher-Giguère et al. (2009).

The initial conditions of the simulation suite are embedded inside periodic cosmological boxes with side length $70.4 - 172\,$Mpc, and were generated at $z \approx 99$ using MUSIC (Hahn & Abel, 2011). Each simulation has 600 snapshots saved down to $z = 0$ with an approximate time spacing of $25\,$Myr. For each simulation we assume flat ΛCDM cosmology with parameters generally consistent with the Planck Collaboration et al. (2020): $h = 0.68 - 0.71$, $\Omega_\Lambda = 0.69 - 0.734$, $\Omega_{\rm m} = 0.266 - 0.31$, $\Omega_{\rm b} = 0.0455 - 0.048$, $\sigma_8 = 0.801 - 0.82$ and $n_s = 0.961 - 0.97$.

4.3.2. Metrics of orbital radius. Throughout this paper we present results in a cylindrical coordinate system. For each galaxy, the axes of are defined via the moment of inertia tensor at a given snapshot. We calculate the moment of inertia tensor using the youngest 25% of all stars which encompass 90% of the total mass within 10 kpc of the center of the galaxy. We calculate the center of the galaxy using an iterative zoom-in method.

Given that stars are not in general on perfectly circular orbits, the definition of the orbital radius of a star is inherently ambiguous. The definition of orbital radius, and the method of calculation, is inconsistent throughout the literature, with different authors using different metrics (for example Halle et al., 2015; El-Badry et al., 2018; Lian et al., 2022; Okalidis et al., 2022). We investigate 5 metrics of orbital radius. Our default metric, $R_{\rm phys}$, is simply the instantaneous radius with respect to the galactic center. We plot this as a thick black line for all figures which show the different radial metrics on the same plot.

In addition, we include the spatially averaged radius of stars $\langle R_{\rm R} \rangle$, which we define as:

$$\langle R_{\rm R} \rangle = \frac{R_{\rm apo} + R_{\rm peri}}{2} \tag{4.1}$$

where $R_{\rm apo}$ is the apocenter of the orbit and $R_{\rm peri}$ is the pericenter. In practice, we calculate this by tracking the physical radius of each star particle at each snapshot and averaging adjacent apocenters and pericenters, calculated using scipy.signal.find_peaks.

We also measure the time-averaged radius, $\langle R_{\rm t} \rangle$, which we define as:

$$\langle R_{\rm t} \rangle = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} R(t)\, dt \tag{4.2}$$

where t_1 and t_2 are the times of adjacent apocenters and pericenters. In practice, we calculate this using numpy.trapz to perform a trapezoidal integral over time between adjacent apocenters and pericenters, and we then divide by the time interval between these radii.

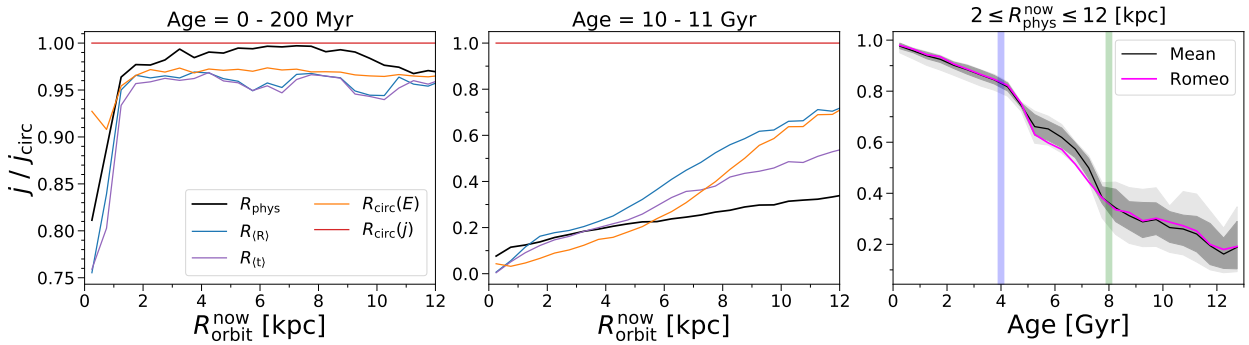


FIGURE 4.2. **Orbital angular momentum of stars today. Left**: The median specific angular momentum of stars younger than 200 Myr at their time of formation, relative to the angular momentum of a circular orbit at their radius, versus current orbital radius $R^{\rm now}_{\rm orbit}$. By definition, $R_{\rm circ}(j)$ is unity. Stars at $R^{\rm now}_{\rm orbit} \lesssim 1.5$ kpc tend to be on less circular, bulge-like orbits. Beyond this radius, this ratio is approximately flat and $\gtrsim 0.95$; stars currently form on nearly circular orbits. **Center**: Same, but measuring orbits today for stars with ages $10-11$ Gyr, which formed mostly during the pre-disk era. This ratio increases with $R^{\rm now}_{\rm phys}$, indicating that old stars are on more circular orbits at larger radii, although the median is always $\lesssim 0.7$. **Right**: The median specific angular momentum of stars relative to that of a circular orbit at the same $R^{\rm now}_{\rm orbit}$ at $z = 0$, versus age, for all stars at $R^{\rm now}_{\rm phys} = 2 - 12$ kpc (excluding the bulge region), averaged across radial bins of width 1 kpc. The black line shows the mean, the dark and light shaded regions show the 68th percentile and full scatter across 11 galaxies. The fuchsia line shows Romeo, whose disk started to form the earliest, ≈ 11 Gyr ago (McCluskey et al., 2023), but which follows a similar trend as the median. The vertical bands separate the pre-disk (age $\gtrsim 8$ Gyr), early-disk (age $\approx 4-8$ Gyr), and late-disk (age $\lesssim 4$ Gyr) eras. Stars that formed in the pre-disk era are generally on highly eccentric orbits today. The ratio increases rapidly in the early-disk era, corresponding to the spin-up of the disk, and it flattens somewhat near unity during the late-disk era, when most stars were born on near-circular orbits.

Following Abadi et al. (2003), we also include the radius of a star on a circular orbit with the same total energy, $R_{\rm circ}(E)$, which we define as:

$$2\Phi(R_{\rm circ}(E)) + R_{\rm circ}(E)\frac{d\Phi(R)}{dR}\bigg|_{R=R_{\rm circ}(E)} = E \tag{4.3}$$

where Φ is the gravitational potential of the disk and E is the sum of the sum of the kinetic and potential energy of a star particle. The simulations self-consistently compute

the potential for each particle using all of the particles and output the potential at each snapshot. We calculate the potential of the disk using the average potential of all star, gas, and dark matter particles contained within ± 0.3 kpc of the galactic midplane, binned radially in 250 pc width bins out to R_{90}^{*} (see Bellardini et al., 2022). We use an azimuthally averaged potential, however we tested our results in the case of an azimuthally varying potential and found variations in our inferred radii on the order of $\lesssim 5\%$. We assume a constant slope for the potential beyond R_{90}^{*}. We require the potential to monotonically increase with radius, in instances where the potential decreases, we set the value of the potential to be equal to the value of the potential in the adjacent radial bin. We numerically differentiate the potential using numpy.gradient. Using Eq. 4.3, we calculate the energy of a circular orbit at different R. We solve for $R_{\rm circ}(E)$ using the numerical solution to the energy of a circular orbit as a function of radius and the sum of the kinetic and potential energy of the star particles.

Our last metric of radius is the radius of a circular orbit with the same angular momentum, $R_{\rm circ}(j)$, which we define via:

$$R_{\rm circ}(j) = j^{2/3} \left[\left. \frac{d\Phi(R)}{dR} \right|_{R=R_{\rm circ}(j)} \right]^{-1/3} \tag{4.4}$$

where j is the specific angular momentum of a star particle, $v_\phi R_{\rm phys}$. We solve for the gradient of the potential following the same procedure as for $R_{\rm circ}(E)$. We then use Eq. 4.4 to calculate angular momentum of a circular orbit at different $R_{\rm phys}$. We use this numerical solution with the instantaneous angular momenta of star particles to find $R_{\rm circ}(j)$ for each star particle.

There are pros and cons to each of the metrics described here. Metrics such as $R_{\rm phys}$, $R_{\rm circ}(E)$, and $R_{\rm circ}(j)$ are instantaneous measurements of stellar radius, so they provide information about rapid perturbations to stellar orbits. Conversely, $\langle R_{\rm R} \rangle$ and $\langle R_{\rm t} \rangle$ are average quantities of stellar orbits and are undefined until a star has completed at least one half of an orbit, so they obfuscate rapid changes to stellar orbits. However, being averaged

quantities they are less sensitive to non-adiabatic changes to the galactic potential. $R_{\rm circ}(E)$, $R_{\rm circ}(j)$, $\langle R_{\rm R} \rangle$, and $\langle R_{\rm t} \rangle$ are relatively difficult to measure for stars, and have a high degree of uncertainty due to the need to accurately measure stellar velocities and accurately model the galactic potential. These metrics are sensitive to changes in orbital properties such as angular momentum or total energy. $R_{\rm phys}$ is relatively straightforward to measure, however it can vary rapidly due to epicyclic motion. We explore these different radial metrics because one might expect different inferences of stellar redistribution depending on the chosen definition of orbital radius.

4.3.3. Transition Epochs. Throughout our analysis we present results for stars of either different ages as a function of radius or for stars at different radii as a function of age. The age bins we present results for are not randomly selected ages, rather ages that roughly correspond to different epochs in the formation history of our sample. Furthermore, for all results presented as a function of age, we show lightly shaded lines demarcating the transition regions to guide the reader's eye.

Observational evidence of high redshift galaxies indicates that they are thicker and clumpier than their lower redshift counterparts (e.g. Glazebrook, 2013; Wisnioski et al., 2015). Surveys exploring the dynamics of galaxies across redshift indicate that galaxies become more rotationally settled with time, i.e. dynamically colder (e.g. Kassin et al., 2012; Übler et al., 2019).

Various analyses of the FIRE-2 simulations have studied different metrics of disk settling across cosmic time. These metrics include analysis of the rotational motion across time (e.g. Ma et al., 2017b), measuring the transition from bursty to smooth star formation (Muratov et al., 2015; Sparre et al., 2017; Yu et al., 2021; Gurvich et al., 2023, e.g.), and measuring the onset of a virialized circumgalactic medium (e.g. Stern et al., 2021).

Following the results of these analyses and work by McCluskey et al. (2023), we explore our results in the context of three different epochs: the 'pre-disk', early-disk', and 'late-disk'

epochs. The precise definitions of these epochs come from the evolution of the ratio of stellar azimuthal velocity to total stellar velocity dispersion (see McCluskey et al., 2023). In general the pre-disk era encompasses lookback times larger than $\approx 8\,\mathrm{Gyr}$, the early-disk era ranges from a lookback time of $\approx 4 - 8\,\mathrm{Gyr}$, and the late-disk era is all lookback times less than $\approx 4\,\mathrm{Gyr}$. During the pre-disk era stars formed on largely isotropic orbits, during the early-disk era galactic disks were still dynamically hot, but stars formed on increasingly circular orbits, and during the late-disk era stars formed on highly rotational orbits.

4.4. Results

4.4.1. Orbital Radii vs Physical Radii. Fig. 4.1 shows the difference between stellar orbital radii and R_{phys} at $z = 0$ for different age stellar populations. For each simulation we calculate the median $R^{\mathrm{now}}_{\mathrm{orbit}} - R^{\mathrm{now}}_{\mathrm{phys}}$ as a function of $R^{\mathrm{now}}_{\mathrm{phys}}$. The shaded regions in Fig. 4.1 show the 68th percentile scatter across the simulation suite. We show the average across the simulations in Fig. 4.1, there are 11 simulations averaged over for $\langle R_{\mathrm{R}} \rangle$ (blue) and $\langle R_{\mathrm{t}} \rangle$ (purple). We only average over 9 simulations for $R_{\mathrm{circ}}(E)$ (orange) and $R_{\mathrm{circ}}(j)$ (red) because the Romulus and Remus simulations do not have stored potential values.

In Fig. 4.1, stars on perfectly circular orbits will show zero difference between their orbital radii and $R^{\mathrm{now}}_{\mathrm{phys}}$. For all metrics of radius, the youngest stars show the most agreement between their orbital radius and $R^{\mathrm{now}}_{\mathrm{phys}}$, however the orbital radii are all systematically smaller at large radii. The discrepancy between orbital radius and $R^{\mathrm{now}}_{\mathrm{phys}}$ typically increases with increasing stellar age and with increasing $R^{\mathrm{now}}_{\mathrm{phys}}$. This is likely correlated with the different epochs of disk formation.

The oldest stars formed in the pre-disk era with no coherent angular momentum and thus the least circular orbits. The intermediate aged stars formed during the early-disk era when disk spin-up was occuring. They generally formed with more coherent angular momentum and thus show more agreement with their physical radius. The youngest stars formed in the

late-disk era, after the disk was well established, they generally formed on the most circular orbits with the most coherent angular momentum (for more discussion see Section. 4.4.2).

$R^{\rm now}_{\rm orbit} - R^{\rm now}_{\rm phys}$ is qualitatively consistent for the different metrics of radii across all ages. However, $R_{\rm circ}(j)$ is systematically smaller than the other metrics of radius at all $R^{\rm now}_{\rm phys}$ and at all ages. The other metrics of radius are generally more quantitatively consistent with one another.

We also measured the distribution of $R^{\rm now}_{\rm orbit} - R^{\rm now}_{\rm phys}$ for all stars in the disk (not shown). The distribution for all metrics of radius are centered at approximately 0 kpc. All metrics are approximately Gaussian, with the exception of $R_{\rm circ}(j)$ which has a negative skew.

4.4.2. Evolution of angular momentum. Figure 4.2 (left) shows the ratio of the median stellar angular momentum to the angular momentum of a circular orbit as a function of $R^{\rm now}_{\rm orbit}$ for stars younger than 200 Myr at their time of formation. By definition $R_{\rm circ}(j)$ is unity.

This ratio is near unity for all stars beyond $R^{\rm now}_{\rm orbit} \approx 1.5$ kpc, indicating young stars are generally forming on circular orbits (e.g. Yu et al., 2021; Hafen et al., 2022). In the inner galaxy the ratio skews lower likely because of bulge kinematics.

The results of McCluskey et al. (2023) indicate stars can experience a large amount of dynamical processing in the first Gyr following their birth. So, we also tested the sensitivity of these results to the size of our age bin by selecting stars with ages ≤ 50 Myr and ≤ 1 Gyr. Selecting stars with ages ≤ 50 Myr leads to generally larger ratios of $j^{\rm form}/j_{\rm circ}$ than selecting stars with ages ≤ 200 Myr and ≤ 1 Gyr. However, the difference for the different metrics of radii are quite small, typically less than ≈ 0.03.

Fig. 4.2 (center) shows the ratio of the median stellar angular momentum to the angular momentum of a circular orbit as a function of $R^{\rm now}_{\rm orbit}$ for stars with ages $10 - 11$ Gyr at $z = 0$. Unlike the youngest stars, this fraction increases with increasing radius. This is likely because older stars at large radii were scattered to larger values of angular momentum.

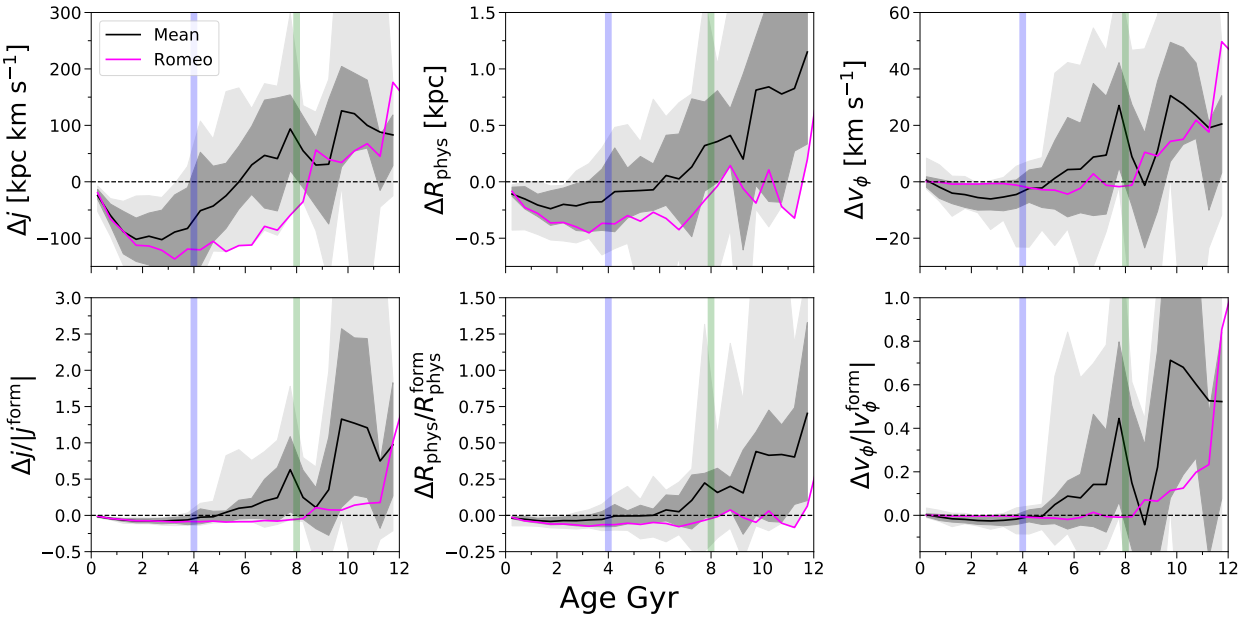


FIGURE 4.3. **Dynamical changes to stars from birth to today across the galaxy.** The change in specific angular momentum j (left), physical radius $R_{\rm phys}$ (middle), and azimuthal velocity v_ϕ (right), for all stars at $R_{\rm phys}^{\rm now} = 2-12\,$kpc (excluding the bulge region). The solid line shows the mean and the dark and light shaded regions show the 68th percentile and full scatter across 11 galaxies. The fuchsia line shows Romeo, which has the earliest-forming disk that started to form $\approx 11\,$Gyr ago. The vertical bands separate the pre-disk, early-disk, and late-disk eras. **Top row**: The median change. Stars that formed during the pre-disk era generally gained j, while stars that formed during the late-disk era generally lost it. Stars that gained j typically did so through an increase in both $R_{\rm phys}$ and v_ϕ, and vice versa. **Bottom row**: Same, but for the median fractional change, relative to the value at formation, which shows more distinctive behavior in each of the eras and diminishes in magnitude over cosmic time, being negative but near zero during the late-disk era. Stars that formed in the pre-disk and early-disk eras generally 'spun up', often significantly, while stars that formed in the late-disk era generally 'spun down', though more modestly. Romeo shows weaker changes at old ages, given its early-forming disk, though it experiences a stronger reduction in j and $R_{\rm phys}$ during the early- and late-disk eras.

Exploring this ratio for other ages shows that, for all radii, this fraction tends to decrease with increasing stellar age. This shows that younger stars tend to be on more circular orbits than older stars. This may largely be driven by younger stars forming on more circular orbits.

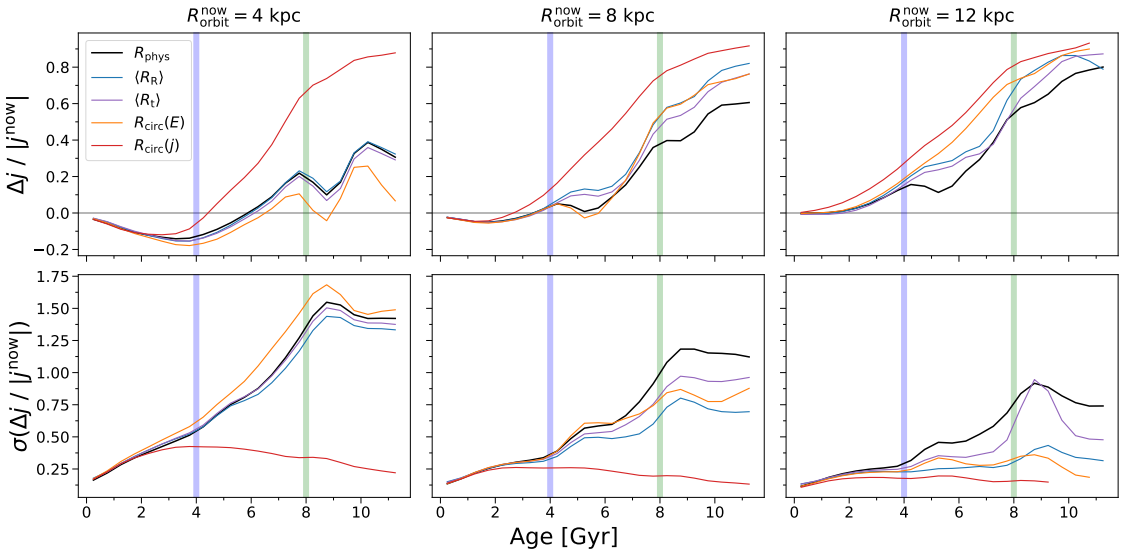


FIGURE 4.4. **Changes to orbital angular momentum of stars from birth to today versus age.** The fractional change in specific angular momentum, relative to its value today, for stars versus age. The columns show stars at different $R_{\rm orbit}^{\rm now}$ (± 0.25 kpc), and for each line we select stars using a different metric for $R_{\rm orbit}^{\rm now}$. The vertical bands separate the pre-disk, early-disk, and late-disk eras. **Top**: The median fractional change. Older stars generally experienced the most significant change, which is systematically positive, and (except for $R_{\rm circ}(j)$) is generally stronger at larger $R_{\rm orbit}^{\rm now}$. The youngest stars generally experienced a modest decrease in j, except at large radii. The transition age, when changes to j transitioned from positive to negative, is younger for stars at smaller $R_{\rm orbit}^{\rm now}$. **Bottom**: Same, but for the half width of the 68th percentile scatter of the fractional change. For all metrics of $R_{\rm orbit}^{\rm now}$ except $R_{\rm circ}(j)$, the scatter increases strongly with age. For stars younger than ≈ 2 Gyr, all metrics of $R_{\rm orbit}^{\rm now}$ yield similar results, but for older stars, measuring orbits using $R_{\rm circ}(j)$ generally leads to a more positive change in the median j (top) but also the smallest scatter in j (bottom).

However, perturbations to the orbits of older stars could also play a role (e.g. McCluskey et al., 2023).

Fig. 4.2 (right) shows the ratio of median stellar angular momentum to the angular momentum of a circular orbit as a function stellar age for stars with $R_{\rm phys}^{\rm now}$ between 2 and 12 kpc. We select a minimum $R_{\rm phys}^{\rm now}$ of 2 kpc because the leftmost panel shows that $j/j_{\rm circ}$ is essentially constant for young stars forming beyond ≈ 2 kpc (however, we explore the impact of our selection radii further in Appendix. C.1). The blue and green lines roughly correspond

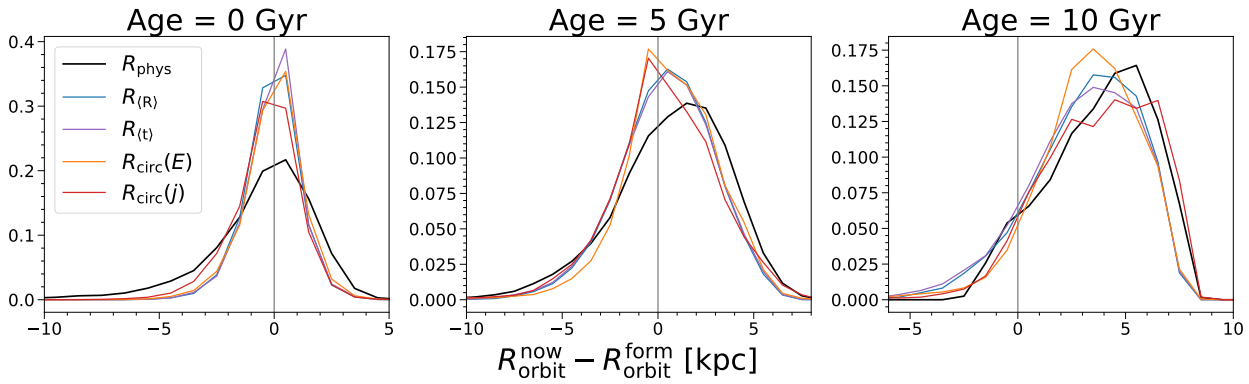


FIGURE 4.5. **Radial redistribution of stellar orbits.** Normalized distribution of the change in $R_{\rm orbit}$ between formation and today, for stars near the Solar annulus today ($R^{\rm now}_{\rm orbit} = 8 \pm 0.25\,{\rm kpc}$), averaged across 11 galaxies (9 for $R_{\rm circ}(E)$ and $R_{\rm circ}(j)$). The panels show stars of different ages today, which, from left to right, formed during the late-disk, early-disk, and pre-disk eras. The different colored lines show results using different metrics for $R_{\rm orbit}$ (at both formation and today), which yield similar results, except using $R_{\rm phys}$ leads to a larger scatter for the youngest stars. The median $\Delta R_{\rm orbit} \approx 0$ for young stars, but older stars experienced more systematic outward redistribution. The width of the distribution is smaller for younger stars and increases to intermediate ages but is essentially unchanged for older stars.

the epochs discussed in Section. 4.3.3. The black line shows the mean across the simulations with the dark shaded region showing the 68th percentile scatter and the light shaded region showing the full range of the distribution. The fuchsia line shows the trends for Romeo, the earliest settling disk in FIRE simulations (McCluskey et al., 2023).

There appear to be transitions in $j/j_{\rm circ}$ that roughly correlate with the transition epochs. Stars which formed prior to 8 Gyr ago tend to have very low angular momentum relative to the angular momentum they would have if they were on a perfectly circular orbit. Stars that formed between 4 and 8 Gyr ago show the steepest increase in $j/j_{\rm circ}$ with age. This early-disk era corresponds to the rapid spin up of the disk and as such stars tend to form on more circular orbits. Stars which formed in the last 4 Gyr show a less steep trend with age. At this point, most of the galaxies have a well defined thin disk, and newly formed stars form primarily on circular orbits.

There is no notable difference between the behavior of Romeo and the behavior of the mean galaxy trend in Fig. 4.2 (right), although Romeo is the earliest settling disk.

Fig. 4.3 (top) shows the median change:

$$\Delta X = X^{\rm now} - X^{\rm form} \tag{4.5}$$

in stellar angular momentum (left), physical radius (center), and azimuthal velocity (right) of the 11 simulations as a function of stellar age. The dark shaded region shows the 68th percentile scatter across the hosts, and the light shaded region shows the full scatter. The fuchsia line shows the trends specifically for Romeo.

Stars which were born prior to $\approx 5.5\,{\rm Gyr}$ ago show a net increase in angular momentum in addition to a net increase in $R_{\rm phys}$ and v_ϕ. The radial trend is partially influenced by our in-situ cut; older stars necessarily formed at smaller radii and are more likely to have scattered to larger radii. The velocity trend is likely a result of disk-spin up; during the pre-disk era stars formed with no coherent rotational velocity, as the disk settled older stars are almost exclusively scattered to larger rotational velocities.

Stars with ages $\lesssim 5.5\,{\rm Gyr}$ generally decreased their angular momentum, $R_{\rm phys}$, and v_ϕ. The onset of a galactic disk likely drives these trends. Stars form on generally circular orbits after the formation of a coherently rotating disk (see Fig. 4.2). Thus, perturbations to stellar orbits will generally drive them away from circular orbits, reducing their angular momentum.

Fig. 4.3 (middle) shows the fractional change in the quantities plotted in the top panel, with respect to their absolute values at formation. There is a clear correlation between the fractional change in these quantities and the different galaxy formation epochs. Stars which formed in the pre-disk era generally have very large fractional changes in j, $R_{\rm phys}$, and v_ϕ. These stars formed with such small initial values of all of these quantities, that even small perturbations lead to large fractional changes. During the early-disk era, as the disk is settling, stars form with increasing angular momentum and increasing v_ϕ, so they generally

experience smaller fractional changes. Finally, in the late-disk era, stars form on nearly circular orbits (Faucher-Giguère, 2018; Hafen et al., 2022; Gurvich et al., 2023; McCluskey et al., 2023) and experience very evolution to their orbits. Additionally, the host-to-host scatter essentially goes away in the late-disk era.

The fractional change in angular momentum of stars in Romeo goes to approximately zero earlier than the simulation average. This, in conjunction with Fig. 4.2 (right) implies that stars generally form on more circular orbits after disk settling. Interestingly, the older stars in the later settling galaxies typically experience increases to their angular momentum that drives them towards orbital circularities comparable to Romeo.

Fig. 4.3 (right) shows a sharp dip to ≈ 0 in both Δv_ϕ and $\Delta v_\phi / |v_\phi^{\rm form}$ at approximately 9 Gyr. This manifests itself in the left column as well, although to a lesser extent. This is likely not a physically meaningful trend. Rather, it highlights that in the pre-disk era there is no well defined disk. We define the cylindrical coordinate system of the galaxy with respect to the moment of inertia tensor of the galaxy. The principal axes of rotation of our galaxies are not reflective of a typical disk galaxy until the disk has settled.

Fig. 4.4 (top) is similar to Fig. 4.3 (left column, bottom row), except it shows the median fractional change in the angular momentum of stars as a function of stellar age for stars currently located at $R_{\rm orbit}^{\rm now} = 4$, 8, and 12 kpc ($\pm 250$ pc). Also, the fractional change is normalized to the present-day angular momentum, rather than the formation angular momentum. So a value of 1 means all of the present-day angular momentum a star has was gained after birth. The different colored lines show the effect of using the different metrics of radius to select stars. The figures are smoothed using scipy.ndimage.gaussian_filter1d with $\sigma = 1.5$.

Independent of radius, older stars typically experienced the larger fractional change to their angular momentum, presumably because they formed with the smaller initial angular

momentum. The youngest stars at each radius typically experienced a net decrease in angular momentum. However, the age marking the transition between stars typically gaining angular momentum and stars typically losing angular momentum is radially dependent. The transition occurs at older ages in the inner disk and younger ages in the outer disk.

For all radial bins, the similarity between the different metrics of radius, seen for young stars, is explained by Fig. 4.1. The difference between the physical radius of stars and the different metrics of radius increases with increasing stellar age. Thus the stellar populations included in the youngest age bins in Fig. 4.4 are almost identical. The older age bins sample different populations of stars, mixing stars of various $R_{\rm phys}$. For the oldest stars there is qualitative agreement between the different metrics of radius. There is quantitative agreement between all metrics of radius too, the exception being stars selected via $R_{\rm circ}(j)$, which show a systematically larger fractional change in angular momentum.

Fig. 4.4 (bottom) shows the $1-\sigma$ scatter of The fractional change in angular momentum as a function of age, for the radial bins specified in the top row. At all radii the scatter is consistent between the different radial metrics for stars younger than $\approx 3\,{\rm Gyr}$. For a given radial region, the scatter generally increases with increasing stellar age, the exception being for stars selected via $R_{\rm circ}(j)$, which show decreasing scatter with increasing stellar age for all stars older than a few Gyr. The scatter is also radially dependent, for a given age stellar population, the scatter is typically largest at the smallest radii.

4.4.3. Distributions of delta R. Fig. 4.5 shows the normalized distribution of $R_{\rm orbit}^{\rm now} - R_{\rm orbit}^{\rm form}$ for 3 different age ($0-1$, $5-6$, and $10-11$ Gyr) stellar populations located in the solar neighborhood ($R_{\rm orbit}^{\rm now} = 8 \pm 0.25\,{\rm kpc}$). The different colored lines represent the distribution, averaged across the 11 hosts (9 hosts for $R_{\rm circ}(j)$ and $R_{\rm circ}(E)$), when defining the present-day radius and the change in radius via different orbital radii.

Fig 4.5 (left) shows the distribution for the youngest subset of stars. The shape of the distribution is qualitatively similar regardless of the chosen metric of radius. Generally,

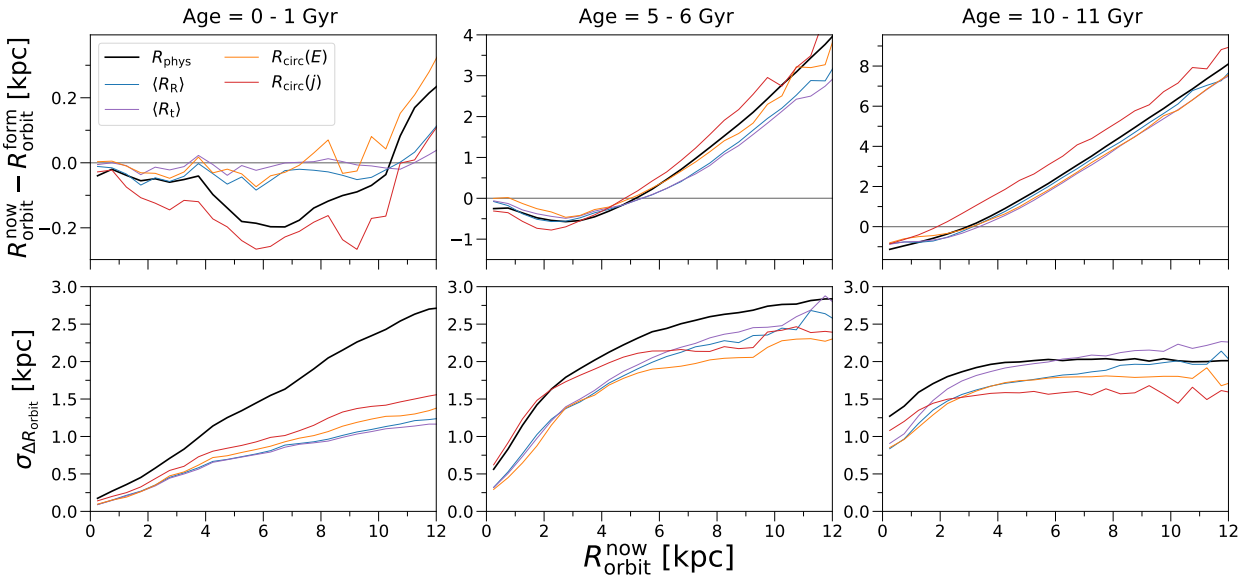


FIGURE 4.6. **Radial redistribution of stellar orbits versus radius.** The change in the $R_{\rm orbit}$ of stars from formation to today versus $R^{\rm now}_{\rm orbit}$. The panels show stars of different ages today, which, from left to right, formed during the late-disk, early-disk, and pre-disk eras. The lines show results using different metrics for $R_{\rm orbit}$ (at both formation and today). **Top**: The median change in the $R_{\rm orbit}$. Similar to Figure 4.7, stars of increasing age experienced increasingly positive redistribution, and young stars generally experienced negative redistribution, with a transition radius that is smaller for older ages. **Bottom**: Same, but for the half width of the 68th percentile scatter. At all ages, the scatter initially increases with $R^{\rm now}_{\rm orbit}$ but tends to saturate, especially for older stars. When self-consistently using the same metric of $R_{\rm orbit}$ to measure redistribution and radius today, the different metrics yield similar results (unlike for using $R^{\rm now}_{\rm phys}$), though the scatter in $R_{\rm phys}$ is generally the largest.

stars younger than 1 Gyr have a peak redistribution of ≈ 0 kpc. The different radial metrics generally show the same scatter (≈ 2.6 kpc), however the scatter in $R_{\rm phys}$ is systematically larger than all other metrics (by a factor of ~ 2. Thus, most of the change in radius the youngest stars have experienced must result from blurring/kinematic heating rather than churning/cold torquing (Sellwood & Binney, 2002; Daniel et al., 2019). This agrees with the results in Fig. 4.4, which shows stars younger than a few Gyr have experienced little to no fractional changes in angular momentum.

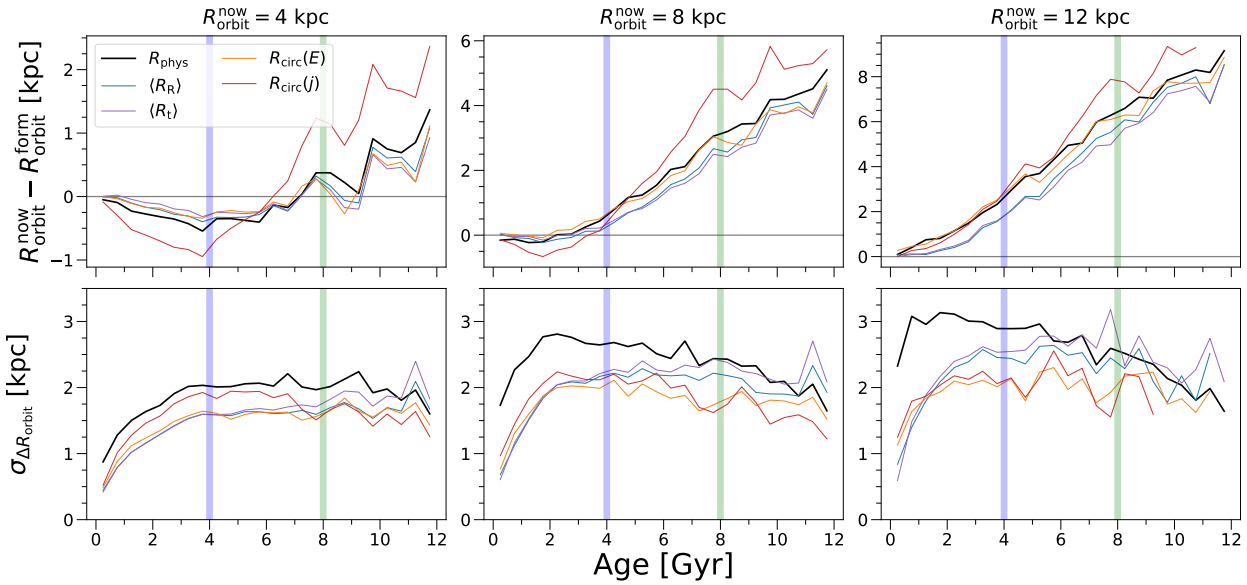


FIGURE 4.7. **Radial redistribution of stellar orbits versus age.** Similar to Figure 4.6, but showing trends versus age in bins of $R^{\rm now}_{\rm orbit}$. The vertical bands separate the pre-disk, early-disk, and late-disk eras. **Top**: The median change in the $R_{\rm orbit}$ of stars from formation to today. In general, $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ becomes increasingly positive with age. The youngest stars generally redistributed slightly inward, while older stars systematically redistributed outward, quite significantly for the oldest stars. This transition occurs earlier for stars at smaller $R^{\rm now}_{\rm orbit}$. **Bottom**: Same, but for the half width of the 68th percentile scatter. The scatter initially increases with age up to ≈ 2 Gyr, after which *the scatter in radial redistribution is approximately constant with age, with a slight decrease*. When self-consistently using the same metric of $R_{\rm orbit}$ to measure redistribution and radius today, the different metrics yield similar results (unlike for using $R^{\rm now}_{\rm phys}$, see Figure C.2), though the scatter in $R_{\rm phys}$ is generally the largest.

Fig 4.5 (middle and right) shows the distribution for older stellar populations. Stars have generally redistributed toward larger radii, and the oldest stars have redistributed the largest amount. This follows in part from our definition of in-situ stars. Stars must have formed within R^{*}_{90}, which decreases with increasing lookback time (see Bellardini et al., 2022). Older in-situ stars necessarily formed in a more compact region and are thus more likely to have redistributed outwards.

We also measure the radial redistribution of stars in different bins of $R_{\rm orbit}^{\rm now}$ (not shown). For the youngest stars, the median of the distribution, averaged across the different radial metrics ranges from $\approx -0.06\,$kpc at $R_{\rm orbit}^{\rm now} = 4\,$kpc to $\approx 0.1\,$kpc at $R_{\rm orbit}^{\rm now} = 12\,$kpc. The width of the distribution shows more dependence on radius, with σ ranging from $\approx 0.8\,$kpc to $\approx 1.7\,$kpc at $R_{\rm orbit}^{\rm now} = 4$ and 12 kpc, respectively.

Although not apparent in the youngest stars, the peak of the distribution exhibits a dependence on $R_{\rm orbit}^{\rm now}$ for older stellar populations. For stars aged $5-6\,$Gyr the median of the distribution is $\approx -0.57\,$kpc at $R_{\rm orbit}^{\rm now} = 4\,$kpc, and it is $\approx 2.1\,$kpc for stars at 12 kpc. This evolution with radius is even more pronounced for stars with ages $10-11\,$Gyr; the median is $\approx 0.12\,$kpc at $R_{\rm orbit}^{\rm now} = 4\,$kpc and $\approx 4.9\,$kpc at $R_{\rm orbit}^{\rm now} = 12\,$kpc. We expect a positive shift in the median of the distribution with increasing radius, because the maximum redistribution stars experience depends on $R_{\rm orbit}^{\rm now}$.

Similar to the trends seen in the young stars, the scatter of the distributions increases with increasing radius for the older stars. For stars with ages $5-10\,$Gyr the scatter increases from $\approx 2.1\,$kpc to $\approx 3.1\,$kpc with increasing radius. For stars with ages $10-11\,$Gyr the scatter increases from $\approx 2\,$kpc to $\approx 3\,$kpc.

4.4.4. Radial redistribution by location. Fig. 4.6 (top) shows the median $R_{\rm orbit}^{\rm now} - R_{\rm orbit}^{\rm form}$ of young, intermediate, and old stars as a function of their present-day orbital radii. We present the stars plotted as a function of their physical radii in Appendix. C.2. There is generally qualitative and quantitative agreement between the different metrics of stellar radius. However, for the oldest stellar population $R_{\rm orbit}^{\rm now} - R_{\rm orbit}^{\rm form}$ of $R_{\rm circ}(j)$ is an outlier. It is systematically more positive than the other orbital metrics.

For all ages, stars presently in the innermost region of galaxies have, on average, redistributed inwards, and stars in the outer regions of galaxies have redistributed outwards. The orbital radius at which stars transition from an average inward redistribution to an average outward redistribution changes with increasing stellar. Stars younger than $\approx 1\,$Gyr

and with $R^{\rm now}_{\rm orbit} \lesssim 11\,{\rm kpc}$ have on average redistributed inward. In contrast, older stars (age $10-11\,{\rm Gyr}$) with $R^{\rm now}_{\rm orbit} \lesssim 3\,{\rm kpc}$ redistributed inward on average. This is in part a result of our in-situ criteria for stars.

While similar for different age stellar populations, the shape of $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ as a function of radius changes with increasing stellar age. For the oldest stars the relationship is nearly linear. For younger stars the relationship is approximately constant with radius, out to $\approx 11\,{\rm kpc}$. Whereas the redistribution of intermediate and older stars typically decreases with increasing $R^{\rm now}_{\rm orbit}$ for stars in the inner disk and increases with increasing $R^{\rm now}_{\rm orbit}$ for stars beyond a few kpc of the galactic center.

The results for intermediate and old age stars make sense given the growth of the galaxy with time. Our in-situ selection criteria necessarily means that for each age there is maximum radius R^{*}_{90} within which all stars must have formed. For all $R^{\rm now}$ beyond this stars necessarily have a positive radial redistribution which goes approximately as $R^{\rm now}_{\rm phys} - R^{*}_{90}$.

Fig. 4.6 (bottom) shows the median $1-\sigma$ scatter in $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ as a function of $R^{\rm now}_{\rm orbit}$. The $1-\sigma$ scatter generally increases with increasing stellar age. The exception is for the oldest stars, where the scatter plateaus at $\approx 4\,{\rm kpc}$. The scatter for the oldest stars is essentially set by R^{*}_{90} of the disk when they formed. Old stars at sufficiently large radii all have nearly the same value of $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ where the only variation is at what radius within R^{*}_{90} they formed. This is also true for the younger stellar populations, but the R^{*}_{90} value of the disk at their time of formation is sufficiently large as to not result in this trend.

Fig. 4.6 (bottom) also shows that for a given stellar age, the $1-\sigma$ scatter in radial redistribution generally increases with increasing present-day radius. This is largest when looking at radial redistribution of $R_{\rm phys}$ and is generally consistent for all other metrics of radius. However, intermediate and old age stars show more similarity between the different metrics of radius than young stars.

In Appendix. C.3 we present a results similar to Fig. 4.6, but for stars binned by their formation radii, rather than their present radii. The trends host averaged results are generally opposite the results presented in Fig. 4.6 (top). The youngest stellar populations generally show positive radial redistribution and older stellar populations generally show negative radial redistribution, except in the inner disk. The reason for the differences is likely due to the negative density profile of the galaxies. Although, most stars born at a given radius systematically redistribute to smaller radii some redistribute to larger radii. Thus, when selecting stars by present-day radii, most stars will have redistributed outwards to that radius rather than inwards because stellar densities are largest at small radii.

4.4.5. Radial redistribution versus age. Fig. 4.7 (top) shows the median change in radius of stars, averaged across the simulations, as a function of age for stars selected in bins of $R^{\rm now}_{\rm orbit}$ (width of 0.5 kpc). To see the stars selected in bins of $R^{\rm now}_{\rm phys}$ see Appendix. C.2. Additionally, we show the stars selected in bins of $R^{\rm form}_{\rm orbit}$ in Appendix: C.3. The vertical shaded bands show the times roughly corresponding to the pre-disk to early-disk transition and the early-disk to late-disk transition.

For all radii, $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ tends to be positive and to increase with increasing age. However, for the youngest stars $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ is negative and decreases with increasing age until ≈ 4 Gyr for $R^{\rm now}_{\rm orbit} = 4$ kpc and until ≈ 2 Gyr for $R^{\rm now}_{\rm orbit} = 8$ kpc. The generally positive redistribution of stars makes sense in the simplified case of Gaussian redistribution of stars. The FIRE galaxies have a negative density gradient (e.g Bellardini et al., 2022), so in the event of random redistribution of stars more stars at a given radius will have redistributed outward than inward.

For all metrics of radius, $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ tends to increase with increasing $R^{\rm now}_{\rm orbit}$. This is in part driven by the asymmetry of the problem. We analyze only stars which formed in-situ, so there is a maximum inward redistribution stars can experience and a maximum outward redistribution, set by $R^{\rm now}$. Stars presently at larger radii necessarily have a larger possible

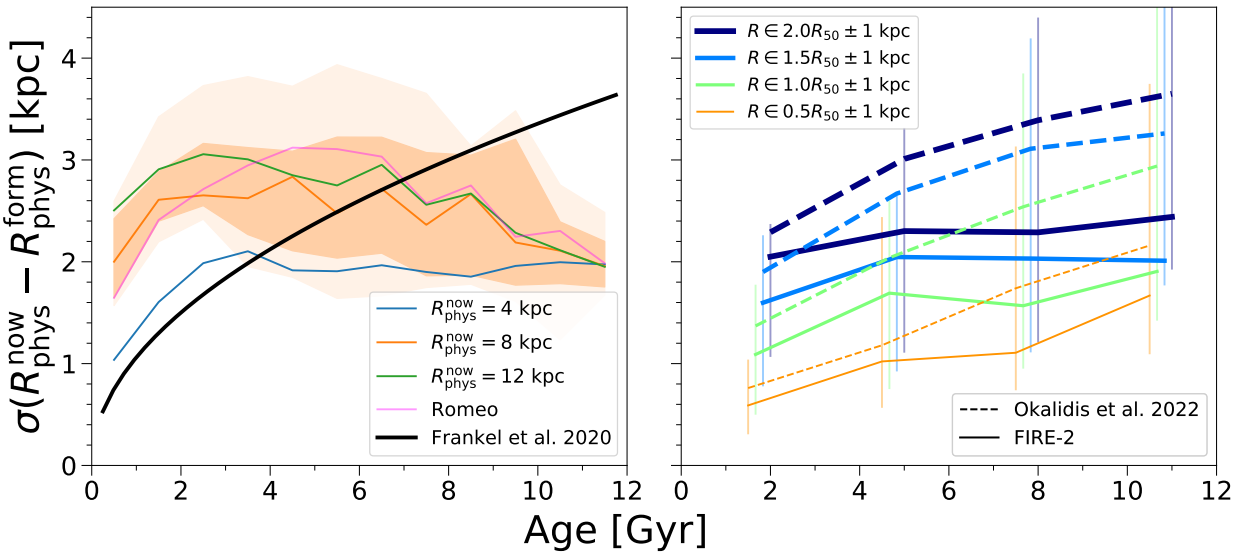


FIGURE 4.8. **Comparing our results with previous works. Left**: The half width of the 68th percent scatter of $R^{\rm now}_{\rm phys} - R^{\rm form}_{\rm phys}$, as a measure of $1-\sigma$, for stars with $R^{\rm now}_{\rm phys} = 4$, 8, and 12 kpc ($\pm 1$ kpc) versus age. The solid lines show the mean at each radius, and the shaded regions show the 68th percentile and full scatter across 11 galaxies (for stars at $R^{\rm now}_{\rm phys} = 8$ kpc). The fuschia line shows Romeo, our earliest-forming disk (≈ 11 Gyr ago), at 8 kpc, which shows a similar trend. The amount of radial redistribution increases moderately with radius, and it increases with age up to $\approx 2-3$ Gyr, but it is flat or decreasing with age for earlier-forming stars. For comparison, the black line shows the (radially independent) model from Frankel et al. (2020). The FIRE-2 simulations show substantially less age dependence: the amount of redistribution is larger for young stars and smaller for older stars. **Right**: Same, but comparing against the results from the Auriga simulations in Okalidis et al. (2022), matching their selection in $R^{\rm now}_{\rm phys}$ via multiples of R^{*}_{50} (see Bellardini et al., 2021, for additional details). Radial redistribution in the FIRE-2 simulations follows similar qualitative trends as Okalidis et al. (2022), though the normalization is lower (less redistribution) and the radial dependence is weaker, especially for older stars.

range of positive $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ than stars at smaller radii. Stars at smaller radii have a larger possible range of negative $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ than stars at larger radii.

The different metrics of radius are generally consistent when the stars are selected in a self-consistent manner, however this is not the case when stars are selected based on $R^{\rm now}_{\rm phys}$,

i.e. the stellar populations plotted for each metric of radius are consistent (see Appendix. C.2). Critically, this implies that the precise definition of radius does not systematically skew conclusions about radial redistribution. The exception to this consistency is the radial redistribution in $R_{\rm circ}(j)$ which skews systematically more negative at young ages and systematically more positive at old ages.

There is no clear correlation between $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ and the transition from the pre-disk to the early-disk era or the transition from the early-disk to late-disk era. However, there is a minor correlation seen at $R_{\rm phys} = 4\,{\rm kpc}$. The age dependence of $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ shifts from a negative slope with increasing stellar age for stars which formed during the late-disk era to a positive slope with increasing stellar age for stars which formed during the earl-disk and pre-disk eras. The reason for this and the lack of this correlation at larger $R_{\rm phys}$ warrants further investigation beyond the scope of this paper.

Fig. 4.7 (bottom) shows the $1-\sigma$ scatter of $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ from the middle panel of Fig. 4.7. Stars younger than $\approx 2\,{\rm Gyr}$ have the smallest scatter in their redistribution. After $\approx 2\,{\rm Gyr}$ the scatter is relatively constant with increasing age, except for $R_{\rm phys}$ which decreases with increasing age. The decrease in scatter with increasing stellar age for $R_{\rm phys}$ follows from our in-situ selection. Older stars necessarily formed in a smaller region of the disk, so they have a smaller range of $R^{\rm form}_{\rm phys}$ from which to redistribute. Indeed, the scatter for the oldest age stars is $\approx 3-4\,{\rm kpc}$ and the average R^{*}_{90} of the simulations at a lookback time of 12 Gyr is $\approx 4\,{\rm kpc}$ (Bellardini et al., 2022). The scatter slightly increases for the different radial metrics with increasing $R^{\rm now}_{\rm orbit}$. However the increase is small, generally less than $\approx 2\,{\rm kpc}$ for all radial metrics compared at $R^{\rm now}_{\rm orbit} = 4\,{\rm kpc}$ and $R^{\rm now}_{\rm orbit} = 12\,{\rm kpc}$.

Following Okalidis et al. (2022), Fig. 4.8 shows the dependence of scatter in stellar redistribution as a function of stellar age, for different radii. We make 2 different cuts to our data to explore the radial redistribution. The different colored lines show our results at different radii, except the fuchsia line which shows our results for Romeo. The dark shaded region and

the light shaded region show the 68th percentile and the full distribution of our simulations at $R_{\rm phys}^{\rm now} = 8\,{\rm kpc}$, respectively. The solid black line shows the radially-independent model of stellar redistribution from Frankel et al. (2020). Fig. 4.8 (left) shows our fiducial cut, i.e. in-situ stars ($R_{\rm phys}^{\rm form}$ less than 30 kpc comoving and vertical height $\leq 3\,{\rm kpc}$. These results are the same as the results presented for $R_{\rm phys}$ in Fig. 4.7 (bottom).

Compared to the model of Frankel et al. (2020), the FIRE simulations show less age dependence in $\sigma(R_{\rm phys}^{\rm now} - R_{\rm phys}^{\rm form})$. For young stellar ages, (age $\lesssim 1.5\,{\rm Gyr}$) the age dependence of our results and the model of Frankel et al. (2020) are qualitatively similar. However, for older ages, the scatter is essentially invariant with age. The absolute $\sigma(R_{\rm phys}^{\rm now} - R_{\rm phys}^{\rm form})$ in the FIRE simulations is larger than that of Frankel et al. (2020) for stars younger than $\approx 6\,{\rm Gyr}$ and smaller for stars older than $\approx 6\,{\rm Gyr}$.

In Fig. 4.8 (left) we show the results for our earliest settling galaxy (Romeo) at $R_{\rm phys}^{\rm now} = 8\,{\rm kpc}$ as its settling time most closely mirrors that of the MW. Even for the case of our earliest settling galaxy we find significant disagreement with the model of Frankel et al. (2020). Romeo shows larger redistribution in young stars and smaller redistribution in old stars. Additionally, Romeo also shows weaker age dependence than the model from Frankel et al. (2020).

Fig. 4.8 (right) shows our data approximately matching the selection criteria from Okalidis et al. (2022). The dashed lines show the values from Okalidis et al. (2022) Fig. 7. The solid lines show the median results from this work, the error bars show $\pm 1\sigma$. We select stars with a vertical height $|Z| < 2\,{\rm kpc}$, a formation radius within the halo, and a circularity $\epsilon > 0.7$. The stars are binned in age bins 3 Gyr in width.

The normalization in the FIRE simulations is slightly smaller than that of Okalidis et al. (2022). The average R_{50} of the simulations we analyze is $\approx 25\%$ smaller than the hosts in Okalidis et al. (2022) ($\approx 3\,{\rm kpc}$ compared to $\approx 4\,{\rm kpc}$). The normalization of the scatter

increases with increasing radius, so this at least partially explains the discrepancies between our results.

Qualitatively, the trends in this work and in Okalidis et al. (2022) are similar, $\sigma(R_{\rm phys}^{\rm now} - R_{\rm phys}^{\rm form})$ tends to increase with increasing stellar age and the normalization increases with increasing radius. However, the age dependence in our results is less pronounced than in Okalidis et al. (2022).

Radial metric	$t_{\rm lb}[(v_\phi/\sigma_{\rm tot})_{\rm form}) > 1\,{\rm Gyr}$				R^*_{90}				M^*_{90}			
	$7 \leq R \leq 9\,{\rm kpc}$		$2 \leq R \leq 12\,{\rm kpc}$		$7 \leq R \leq 9\,{\rm kpc}$		$2 \leq R \leq 12\,{\rm kpc}$		$7 \leq R \leq 9\,{\rm kpc}$		$2 \leq R \leq 12\,{\rm kpc}$	
	ρ	p-val	ρ	p-val	ρ	p-val	ρ	p-val	ρ	p-val	ρ	p-val
$\Delta\, R_{\rm phys}$	-0.69	0.02	-0.53	0.09	-0.71	0.02	-0.76	0.01	-0.61	0.05	-0.67	0.02
$\Delta\, \langle R_{\rm R} \rangle$	-0.68	0.02	-0.47	0.15	-0.61	0.05	-0.79	< 0.01	-0.86	0.02	-0.63	0.04
$\Delta\, \langle R_{\rm t} \rangle$	-0.77	0.01	-0.41	0.22	-0.65	0.03	-0.75	0.01	-0.76	0.01	-0.56	0.08
$\Delta\, R_{\rm circ}(E)$	-0.71	0.03	-0.20	0.60	-0.59	0.09	-0.65	0.06	-0.80	0.01	-0.37	0.33
$\Delta\, R_{\rm circ}(j)$	0.01	0.98	-0.63	0.07	-0.26	0.50	-0.79	0.01	0.08	0.83	-0.80	0.01
$\sigma(\Delta\, R_{\rm phys})$	0.11	0.75	0.14	0.68	0.82	< 0.01	0.64	0.04	0.26	0.45	0.68	0.02
$\sigma(\Delta\, \langle R_{\rm R} \rangle)$	0.31	0.36	0.18	0.61	0.69	0.02	0.55	0.08	0.51	0.11	0.71	0.02
$\sigma(\Delta\, \langle R_{\rm t} \rangle)$	0.15	0.66	0.18	0.61	0.59	0.06	0.55	0.08	0.55	0.08	0.71	0.02
$\sigma(\Delta\, R_{\rm circ}(E))$	0.42	0.26	0.22	0.57	0.68	0.05	0.64	0.06	0.23	0.55	0.55	0.13
$\sigma(\Delta\, R_{\rm circ}(j))$	0.52	0.15	0.56	0.12	0.90	< 0.01	0.79	0.01	0.48	0.19	0.70	0.04

TABLE 4.1. The correlation strength, ρ, and significance (p-value), between mean $\Delta R_{\rm orbit}$ or the half width of the 68th percentile scatter in $\Delta R_{\rm orbit}$ against: the lookback time to when galaxies transitioned to the early-disk epoch, $t_{\rm lb}[(v_\phi/\sigma_{\rm tot})_{\rm form} > 1]$; the radius encompassing 90% of the stellar mass today, R^*_{90}; and 90% of the stellar mass today, M^*_{90}. For each correlation, we use stars in the solar neighborhood ($R^{\rm now}_{\rm orbit} = 7 - 9\,{\rm kpc}$) and stars across the disk ($R^{\rm now}_{\rm orbit} = 2 - 12\,{\rm kpc}$). The first column lists the metric of orbital radius we use in measuring radial redistribution from birth to today. Generally, both the mean radial redistribution and the half width of the 68th percentile scatter for stars in the solar neighborhood and in the whole disk show statistically significant correlations (p-value < 0.05) with R^*_{90} and M^*_{90}. The mean radial redistribution of stars correlates negatively with galaxy size and mass, but the half width of the 68th percentile scatter in the stellar radial redistribution correlates positively with them. The mean radial redistribution of stars in the solar neighborhood correlate negatively with the transition lookback time and show the only statistically significant trends with the transition lookback time.

4.4.6. Radial redistribution correlations. In this section we test for correlations between the strength of radial redistribution and galaxy properties. Correlations with different galaxy properties inform us about potential mechanisms driving radial redistribution as well as helping to constrain radial redistribution in the MW.

We measure Spearman rank correlations of the mean radial redistribution of all stars with an orbital radius currently in the solar neighborhood ($7 \leq R_{\rm orbit}^{\rm now} \leq 9\,{\rm kpc}$ and all stars in the disk ($2 \leq R_{\rm orbit}^{\rm now} \leq 12\,{\rm kpc}$) with the transition lookback time from McCluskey et al. (2023), when galaxies transition from the pre-disk era to the early disk era ($t_{\rm lb}[(v_\phi/\sigma_{\rm tot})_{\rm form} > 1]$), as well as the galaxy size (R_{90}^*) and mass (M_{90}^*) from Bellardini et al. (2021) . We also measure correlations against the same variables for the half widht of the 68th percentile scatter of the radial redistribution of the stars. We measure correlations for each metric of radius and present them in Table. 4.1. Across all correlations we explore the most statistically significant trends observed are correlations of $\sigma(\Delta\, R_{\rm circ}(j))$ and $\sigma(\Delta\, R_{\rm phys})$ with galaxy size for stars in the solar neighborhood and mean $\Delta\, \langle R_{\rm R} \rangle$ with galaxy size for all stars in the disk.

In general, the mean radial redistribution of stars correlates significantly (p-value < 0.05) with the the transition from the pre-disk to early-disk era, with $R_{\rm circ}(j)$ being the exception. The correlation coefficient is negative, meaning galaxies which transitioned to a disk later show stronger radial redistribution than galaxies which transitioned earlier. A potential cause of this is the breathing mode undergone by galaxies, in which stars are born on outflows due to strong feedback events (e.g. El-Badry et al., 2016; Yu et al., 2020). This is further supported by the lack of correlation seen for $R_{\rm circ}(j)$, as the breathing modes heat stellar orbits rather than torquing them.

Interestingly there is also a significant negative correlation between the mean radial redistribution and the galaxy size. The exception being for both $R_{\rm circ}(E)$ and $R_{\rm circ}(j)$. Smaller galaxies tend to show statistically larger, although still relatively small radial redistribution,

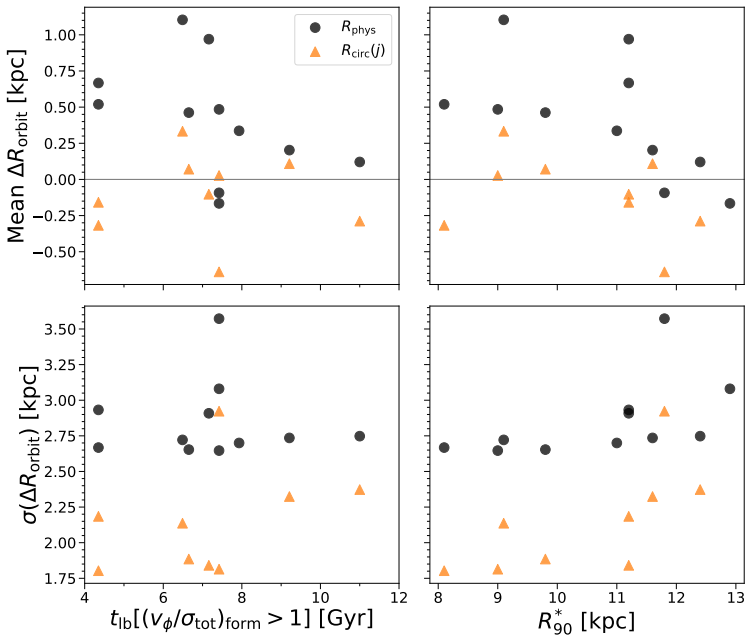


FIGURE 4.9. **Top**: Mean radial redistribution of all stars with $R^{now}_{orbit} = 7 - 9$ kpc as a function of the lookback time of disk onset (from McCluskey et al., 2023), $t_{lb}[(v_\phi/\sigma_{tot})_{form} > 1]$ (left), and the radius that encompasses 90% of the stellar mass, R^*_{90} (as in Bellardini et al., 2021). Table 4.1 shows the correlations for all metrics of orbital radius; here we show only R_{phys} and $R_{circ}(j)$. Correlations with R_{phys} are generally representative of correlations with all other radial metrics except $R_{circ}(j)$. For all radial metrics, except $R_{circ}(j)$, there is a statistically significant negative correlation between the mean ΔR_{orbit} and $t_{lb}[(v_\phi/\sigma_{tot})_{form} > 1]$; galaxies that settled later show stronger outward redistribution. There is also a negative correlation between the mean ΔR_{orbit} and R^*_{90} (except for $R_{circ}(j)$), so smaller galaxies have experienced more positive radial redistribution than larger galaxies. **Bottom**: Similar, except for the half width of the 68th percentile of radial redistribution. Now, in general, there are no significant correlations with when the disk started to settle and the amount of radial redistribution. However, there are statistically significant positive correlations between the $\sigma(\Delta R_{orbit})$ and R^*_{90}, larger galaxies have slightly broader distributions of ΔR_{orbit}.

in the solar neighborhood than larger galaxies. This same trend is also true for correlations with galaxy stellar mass. Stars presently in the solar neighborhood in less massive galaxies tend to have experienced more positive radial redistribution than the stars in less massive galaxies. However the mechanism driving this is beyond the scope of this paper, and warrants future investigation.

Table. 4.1 shows that regardless of the metric of radius chosen, for stars in the solar neighborhood, there is no significant correlations between the 68th percentile scatter in $\Delta R_{\rm orbit}$ and the transition from the pre-disk to early-disk era. This is also true for correlations with galaxy stellar mass. However, there is a statistically significant correlation between the scatter in $\Delta R_{\rm orbit}$ and the galaxy size. Larger galaxies show a larger spread in radial redistribution for all metrics of radius, aside from $R_{\rm circ}(E)$.

In addition to correlations for stars in the solar neighborhood, Table. 4.1, tests the sensitivity of our results to our radial selection by showing trends for stars across the entire disk ($2 \leq R_{\rm orbit}^{\rm now} \leq 12\,{\rm kpc}$). Generally, for all metrics of radius, the correlations with the transition lookback time tend to decrease in statistical significance when selecting all stars rather than just stars in the solar neighborhood. Trends between mean radial change and galaxy size increase in statistical significance when selecting all stars rather than just solar neighborhood stars. However, trends between the 68th percentile scatter and galaxy size decrease in statistical significance when selecting all disk stars. Correlations between mean radial redistribution and galaxy stellar mass are generally slightly less statistically significant when selecting all stars, and the correlations with the 68th percentile scatter are generally more statistically significant when selecting all stars. The radial metric which exhibits the most dependence on our selection criteria is $R_{\rm circ}(j)$. The mean radial change was uncorrelated with galaxy size and mass for stars in the solar neighborhood, but is strongly correlated for all disk stars.

Additionally, although not shown here, we test correlations of mean $\Delta R_{\rm orbit}/R_{90}^{*}$ and $\sigma(\Delta R_{\rm orbit}/R_{90}^{*})$ with R_{90}^{*}. The correlations between find the correlations $\Delta R_{\rm orbit}/R_{90}^{*}$ and galaxy size are essentially the same as the correlations between the mean radial change and R_{90}^{*}. The correlations between $\sigma(\Delta R_{\rm orbit}/R_{90}^{*})$ and galaxy size tend to be weaker than the correlations between $\sigma(\Delta R_{\rm orbit})$ and galaxy size.

We show a representative sample of the correlations in Fig. 4.9. The top row shows mean $\Delta R_{\rm orbit}$ as a function of $t_{\rm lb}[(v_\phi/\sigma_{\rm tot})_{\rm form} > 1]$ and R^*_{90}. The bottom row shows $\sigma(\Delta R_{\rm orbit})$ vs $t_{\rm lb}[(v_\phi/\sigma_{\rm tot})_{\rm form} > 1]$ and R^*_{90}. The circles show trends for $R_{\rm phys}$ and the triangles for $R_{\rm circ}(j)$.

4.5. Summary and Discussion

4.5.1. Summary. In this paper we characterized the radial redistribution of stars in 11 MW/M31-mass cosmological zoom-in simulations from the FIRE-2 suite. We explore the results defining stellar radius using 5 different metrics: the instantaneous position of the star, 2 radii based on the galactic potential, and 2 radii based on orbit tracking. We measured the typical change in angular momentum of stars as a function of stellar age. We also subdivided stars based on age and measured the radial redistribution as a function of present-day and birth radius, and we subdivided stars into different present-day radii and measured radial redistribution as a function of stellar age.

Our key results are as follows:

- *Stars prior to the late-disk era typically gain angular momentum and stars born during the late-disk era typically lose angular momentum* (see Figs. 4.3 and 4.4). Both old and young stars experience similar absolute changes to their angular momentum, but older stars experience larger fractional changes to their angular momentum.
- *Stars currently in the outer disk experienced more radial redistribution than stars in the inner disk* (see Figs. 4.7 and 4.6). Stars currently in the inner few kpc of the disk have experienced near 0 kpc of radial redistribution on average. Stars presently in the outermost region of the galaxy typically experienced large positive radial redistribution.
- *The oldest stars have experienced the most significant radial redistribution* (see Figs. 4.6 and 4.7). Across nearly all radii, the youngest stars experienced little to no radial redistribution $\lesssim 1$ kpc. Intermediate and old stars in the inner disk

experienced radial redistribution that scales nearly linearly with $R^{\rm now}$ because the oldest stars formed in the central few kpc of the disk.

- *The 68th percentile of* $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ *is radially dependent* (see Fig. 4.8. A simplified model in which $\sigma(\, R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit})$ is radially independent does not accurately reflect stellar radial redistribution. Especially for younger stars, the 68th percentile scatter is smaller at smaller radii than at larger radii.
- *The mean radial redistribution and the scatter in radial redistribution correlated with galaxy properties* (see Table. 4.1). Galaxies which established a disk earlier show less positive mean radial redistribution than galaxies which formed a disk later. However, $\sigma(\Delta R_{\rm orbit})$ shows no significant correlation with when the galactic disk settled. The mean radial redistribution in larger galaxies is less positive than in smaller galaxies, however $\sigma(\Delta R_{\rm orbit})$ is larger for larger galaxies.
- *Self-consistency in radial redistribution analysis yields consistent results across all metrics of radius* (see Figs. 4.5, 4.6, and 4.7). The scale of stellar redistribution is independent of definition of stellar radius, whether that be some measure of the guiding center radius or the physical radius, as long as the $R^{\rm now}$ and $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ are measured self-consistently.

4.5.2. Discussion. In this paper we explore radial redistribution of stars in 11 simulated galaxies taken from the *Latte* (Wetzel et al., 2016) and ELVIS (Garrison-Kimmel et al., 2019a,b) suites. We compare these results for 5 different methods of defining radius: $R_{\rm phys}$, $\langle R_{\rm R} \rangle$, $\langle R_{\rm t} \rangle$, $R_{\rm circ}(E)$, and $R_{\rm circ}(j)$ (defined in Section. 4.3.2).

In Section. 4.4.1 we show that, young stars are generally on circular orbits. This agrees with results from Yu et al. (2021, 2022); Hafen et al. (2022) which show that, after disk settling, stars form in a thin-disk on circular orbits. Our results imply that stars do not form on circular orbits throughout the lifetime of a galaxy.

Older stars are generally on less circular orbits than young stars, in agreement with observations (e.g. Dehnen & Binney, 1998). Importantly, this is not a result of older stars forming on circular orbits and being scattered to non-circular orbits. As shown in Section. 4.4.2, older stars typically experience a 'spin-up' in their angular momentum. Thus, the lower circularity of the orbits of older stars reflects the birth conditions of these stars rather than dynamical heating of their orbits. This agrees with other analyses from FIRE (e.g. Yu et al., 2022; McCluskey et al., 2023) as well as other simulations (e.g. Pillepich et al., 2019) and observations (e.g. Wisnioski et al., 2015; Swinbank et al., 2017).

Given that Romeo is the earliest settling disk in the suite of simulations we explore (McCluskey et al., 2023), in Section. 4.4.2 we compare trends in Romeo to the average trends in our simulations. The median fractional change in angular momentum as a function of stellar age is approximately 0 for Romeo for all stars which formed after the onset of the early-disk era ($\approx 11\,\mathrm{Gyr}$ ago). Additionally, the scatter in the fractional change to angular momentum, radius, and circular velocity is smallest for stars which formed in Romeo for nearly all ages. This supports our conclusion that stars which form in a rotationally supported disk generally experience fractionally less significant changes to their angular momentum, radius, and circular velocity.

Although stars which form in rotationally supported disks generally experience minimal fractional changes to their angular momentum, stars which formed prior to disk settling experienced large fractional changes to angular momentum. This contrasts with assumptions of Hu et al. (2023) based on Sanders & Binney (2016). Under the assumption of radial migration, the median angular momentum of stars born at a given radius will remain unchanged while the distribution broadens. Figure. 4.4 shows that the median angular momentum of stellar populations can vary drastically between birth and present-day for sufficiently old stellar population, and that the dependence varies on the present-day radial location of

stars. This suggests there are other important dynamical processes which change the orbits of stars.

Similar to the results of Okalidis et al. (2022), we find that the youngest stars in FIRE, regardless of their $z = 0$ radius have generally experienced the smallest redistribution (top row of Fig. 4.5) and the oldest stars have experienced the most radial redistribution (bottom row of Fig. 4.5). Additionally, the width of the distribution increases with increasing age for a given radius, and the width of the distribution increases with increasing radius for a given age.

Further analysis of the of radial redistribution (see Fig. 4.7) shows that the radial redistribution experienced by stars is strongly correlated with their present-day radius. Stars presently at large radii, regardless of the chosen metric of radius, have typically increased their radius relative to their birth radius, which agrees with the analysis of Halle et al. (2015). Additionally, as with Halle et al. (2015), stars that are in the inner disk have typically redistributed inwards. Interestingly Halle et al. (2015) find that the number of extreme outward migrators ($R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit} > 6\,{\rm kpc}$) decreases in the outer disk when excluding blurring. This is consistent with our findings that $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ is systematically largest for $R_{\rm phys}$ and smaller for all metrics of orbital radii.

In Section. 4.4.5 we explore radial redistribution of stars as a function of stellar age. Specifically in Fig. 4.8 we compare 68th percentile scatter in $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ for stars at different radii as a function of age to results presented in Okalidis et al. (2022) and the model from Frankel et al. (2020). The model proposed by Frankel et al. (2020) suggests no radial dependence to this scatter. However, in agreement with the results of Auriga (Okalidis et al., 2022), we find there is a radial dependence.

For our fiducial selection criteria, binning stars by $R_{\rm phys}$, there is a minor radial dependence to $\sigma(\, R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit})$ which is most extreme in the youngest stars. For no radius do we find consistency between measured scatter in radial redistribution of the oldest stars in

our simulations and the model of Frankel et al. (2020). Even for our earliest settling disk, Romeo, the scatter is systematically smaller in FIRE at large ages and systematically larger for small ages.

The scatter measured in FIRE is also systematically lower than the scatter measured in Okalidis et al. (2022), however this is true across all ages and across all radii. A potential cause of this is that the Auriga simulations have systematically larger disks than the FIRE simulations, and there is a radial dependence to the scatter. So, for a given fraction of R_{50} the selection radius in FIRE is systematically smaller than in Auriga.

In Section. 4.4.6, we look at correlations of the mean radial redistribution of stars and of the half width of the 68th percentile scatter in the radial redistribution of stars with the time at which galaxies transitioned from the pre-disk to the early-disk era and the size of the galaxies. Surprisingly, there is no significant correlation between the disk settling time and $\sigma(\Delta R_{\rm orbit})$. This suggests that the primary driver of radial redistribution of stars in the FIRE simulations is not radial migration/resonant scattering. That mechanism of radial redistribution only occurs in galaxies with established disks and non-axisymmetric structures like spiral arms. Thus, if it were the dominant mechanism of radial redistribution one would expect a strong correlation with disk settling time. Stars in larger galaxies tend to experience less positive radial redistribution than stars in smaller galaxies, which is potentially a result of strong feedback episodes leading to stars being born on outflowing gas (e.g. El-Badry et al., 2016; Yu et al., 2020). However, this trend warrants future investigation.

Additionally, the mean radial redistribution in galaxies correlates negatively with the lookback time at which galaxies transitioned from the pre-disk era to the early-disk era. This hints at the influence of different mechanisms of radial redistribution in different epochs. It further supports the idea that in the pre-disk era stars are being born on outflows. Also, it implies that the mechanisms driving radial redistribution in the early- and late-disk eras do

not preferentially redistribute stars inwards or outwards. The galaxies that have had a disk the longest show the least amount of radial redistribution, on average.

Perhaps the most important finding in this paper is the quantitative consistency of stellar radial redistribution between all metrics of radii, when performing analysis in a self-consistent manner. This is emphasized in Figs. 4.5, 4.6, and 4.7. In other words, when measuring the radial redistribution on stars, the result is agnostic to the definition of stellar radius. This suggests, for studies of radial redistribution a preference should be given to whatever method of measuring stellar radius is easiest and most accurate.

Acknowledgements

We performed this work using the GizmoAnalysis package (Wetzel & Garrison-Kimmel, 2020), the Astropy package (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018), as well as libraries from Numpy (Harris et al., 2020), SciPy (Virtanen et al., 2020), and Matplotlib (Hunter, 2007). We also made use of the online tool WebPlotDigitizer (Rohatgi, 2022).

Data Availability

The data in these figures is available at `https://mbellardini.github.io/`. The FIRE-2 simulations are publicly available (Wetzel et al., 2023) at `http://flathub.flatironinstitute.org/fire`. Additional FIRE simulation data is available at `https://fire.northwestern.edu/data`. A public version of the GIZMO code is available at `http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html`.

CHAPTER 5

Conclusion

5.1. Summary

In this dissertation I made use of 11 MW/M31 mass hydrodynamic cosmological zoom-in simulations from the *Latte* and 'Elvis' simulation suites of FIRE to explore the feasibility of chemically-tagging MW stars. I first used these simulations to characterize and quantify the scale of elemental abundance homogeneity in galaxies across cosmic time. To help with chemically tagging stars born in more homogeneous environments, I then quantified the radial redistribution of stellar populations in these simulations at redshift 0. I summarize the main results and the implications of each chapter in the subsequent paragraphs.

Chapter 2 is one of the first studies of cosmological simulations to characterize the full 3-D evolution of gas-phase elemental abundances. Given the relative difficulty of taking spatially resolved measurements of high redshift galaxies, simulations are currently one of the best methods to characterize the spatial distribution of elemental abundances in early universe galaxies. In fact, the question of the time evolution of galactic radial abundance gradients is still contested, with some groups claiming a shallowing of the profile with time, other groups claiming a steepening, and some groups claiming a steepening followed by a shallowing. I found that galaxies generally show disks which are well mixed radially at high redshifts, in fact the dominant source of elemental abundance variation comes from azimuthal abundance fluctuations. This complicates the work of chemically tagging stars, as stellar birth locations are not uniquely defined in elemental abundance space. This chapter also presented results which challenged a standard assumption of galactic chemical evolution models: elemental abundances are well mixed at a given radius across time, such that all

stars born at a given location will have roughly the same elemental abundances. Crucially, in this chapter I presented a characterization of the transition lookback time at which simple one dimensional chemical evolution models breakdown, essentially the time at which the main source of elemental abundance variations in galaxies shifts from azimuthal to radial variations.

Chapter 3 expands upon the results of Chapter 2 by quantifying the 3-D evolution of elemental abundances in newly formed stars, which does not necessarily trace the trends see in all gas. However, the evolution of radial and azimuthal fluctuations of stellar elemental abundances in galaxies generally follows the trends of the gas. One crucial result presented in Chapter 3 is the scatter in [Fe/H] as a function of stellar age, which reflects the competition between azimuthal scatter and radial gradients across the history of a galaxy. The age at which scatter is minimized roughly correlates with the time at which galaxies transitioned from abundance variations being dominated by azimuthal fluctuations to being dominated by radial variations. The most important results presented in Chapter 3 are the functional forms and best fits to the average stellar abundance, radial abundance gradient, and azimuthal scatter of stellar abundances as a function of cosmic time. The fits to the spatial distribution of elemental abundances are invaluable to researchers working on chemical tagging as they are in essence a cosmologically motivated galactic chemical evolution model. Additionally, with the advent of newer telescopes and observational surveys focusing on high redshift galaxies these functional forms provide a prediction for the expected observed spatial distribution of elemental abundances. Whether there is or is not agreement between these predictions and observations the field of galactic archaeology will improve. Agreement with these predictions would further validate the physics implementation in cosmological simulations and increase their usefulness as tools for galactic archaeologists. Disagreement between these predictions and observations will spur an improvement in cosmological simulations.

Chapter 4 explores the typical scale of radial redistribution of stars. This is especially important to quantify for older populations of stars as I previously laid out the relative homogeneity of elemental abundances in the galactic disk at early times. Careful quantification of the radial redistribution of stars can break degeneracies of birth locations for older stars, strengthening the predictive power of chemical tagging. I find that the amount of radial redistribution stars have experienced depends strongly both on their present-day radius and their age. This is in agreement with the results presented in other simulation groups, but contrasts with the observation-based models currently being used by some theorists. Crucially, in Chapter 4, I also explored the different metrics by which observers and theorists typically define the radius of a star. I found that regardless of the definition of radius used, as long as stars are radially binned using the same definition, there are no appreciable differences in the radial redistribution of stars as a function of present-day radius, formation radius, and age. This is the first robust comparison of these different definitions of radius. Thus observers and theorists can freely use their preferred method of measuring stellar radii when chemically tagging stars and reach similar conclusions.

5.2. Future Work

The work I did in this dissertation has set the stage for a multitude of future projects to further explore the viability of chemically tagging stars. Most recently Graf et al. (in preparation) has characterized the present-day 3D spatial distribution of stellar elemental abundances as a function of stellar age. They compare the strength of radial gradients, vertical gradients, and azimuthal abundance of stars at formation and at $z = 0$ in the FIRE simulations to quantify the impact of stellar radial redistribution on the distribution of elemental abundances at present-day. While I quantified the evolutionary history of several MW-mass galaxies, this work allows us to better understand the specific history of the MW.

Additionally, future work should look to quantify the origins of the radial redistribution I characterized in Chapter 4. Understanding the extent to which stars move is important for

chemically tagging stars, but it is also important to quantify the role of different physical drivers of stellar radial redistribution. A more in depth analysis of radial redistribution in cosmological simulations can be done by simply tracking the time evolution of stellar angular momentum, orbital energy, orbital eccentricity, and radius across time. Blurring/kinematic heating of stars would show that stellar guiding radii, defined by their angular momenta, are essentially constant with time while their guiding radius defined via their orbital energy grows with time and their physical radius oscillates with time. Radial migration or churning of stars would show rapid changes to the guiding radii of stars defined by their angular momenta as they experience corotation resonance with galactic bars or spiral arms or radial redistribution due to Lindblad resonances (e.g. Minchev & Famaey, 2010; Minchev et al., 2011, 2012). Characterization of stellar kinematics across time is currently being explored by McCluskey et al. (in preparation).

Yet another direction in which future work can go is understanding the origin of the spatial abundance variations observed in galaxies, particularly azimuthal abundance fluctuations. This has already been explored in part by Khoperskov et al. (2023); Orr et al. (2023). They have found that azimuthal elemental abundance fluctuations are intrinsically linked to spiral structure in galaxies. Using hydrodynamic simulations, both groups showed that the azimuthal abundance fluctuations seen in disk galaxies as present-day appear to be a result of gas funneling along spiral structure in galaxies with pre-existing radial abundance gradients. In effect metal rich gas is moved to larger radii and metal poor gas is moved to smaller radii inducing abundance fluctuations at a given radius. Khoperskov et al. (2023); Orr et al. (2023) also find that local enrichment alone is insufficient to produce the abundance fluctuations observed in galaxies at present-day. Graf et al. (in preparation) also test correlations between the strength of azimuthal abundance fluctuations and the strength of radial gradients for young stars at present-day in the FIRE simulations, finding that galaxies with

the strongest radial gradients show the largest azimuthal fluctuations. This further supports the notion of radial redistribution driving present-day azimuthal abundance fluctuations.

However, the mechanism responsible for azimuthal abundance fluctuations across cosmic time is not well understood. The work of Khoperskov et al. (2023); Orr et al. (2023); Graf et al. (in preparation) provide a reasonable explanation for the presence of azimuthal fluctuations at $z = 0$, but azimuthal fluctuations induced via radial redistribution in the presence of a pre-existing radial gradient does not explain the high redshift azimuthal abundance fluctuations we present in Chapters 2 and 3. At large redshift galaxies exhibit essentially no radial gradient yet large variations in azimuthal scatter. A careful analysis of the merger histories, local enrichment, and mixing timescales at these redshifts must be done to better understand the physical origin of these fluctuations.

Another important future step is building a model to apply chemical tagging to stars in the FIRE simulations. This model should incorporate all the information from this dissertation on the time evolution of the spatial distribution of elemental abundances as well as the typical radial redistribution experienced by stars outlined in Chapter 4. Assuming the accuracy of the FIRE simulations, this is a means to test the best case scenario for chemically tagging stars. Using an accurate model for the scale of radial redistribution of stars and galactic chemical evolution, we can essentially place an upper limit on the expected precision of chemical tagging.

Galactic archaeology has quickly progressed in popularity, leading to entire surveys dedicated to better understanding the formation and evolution of the MW, like the GALactic Archaeology with HERMES survey. Knowing the discriminating power of chemical tagging is crucial to surveys like this, as it plays an important role in their selection function and their desired measurement precision of elemental abundances. As such, predictions of the spatial distributions of elemental abundances from cosmological simulations is crucial to our understanding of the homogeneity of star forming sites which feeds into galactic chemical

evolution models. Despite the massive investment into observations, spatially resolved elemental abundance maps of high redshift galaxies is still exceedingly difficult. As such, at present, our best understanding of the initial conditions from which stars form at high redshift must come from simulations. However, it is important to continually benchmark simulations against future observations, as observations are the ground truth of the universe. Crucially, simulations must also provide predictions of the time evolution of the spatial distribution of elemental abundances, so that they can continually be tested against constantly improving observations. While we may not be able to exactly simulate the conditions of the proto-MW, continually expanding the number of high-resolution simulations sampling a diverse range of evolutionary histories provides us with the best chance of understanding the evolutionary history of our own galaxy.

APPENDIX A

Chapter 2 Appendices

A.1. Scaled radial profiles

Fig. A.1 compares the host-to-host scatter in radial gradients of [O/H] in gas in our simulated galaxies when scaling these gradients to various galaxy scale radii at $z = 0$. We scale each galaxy's profile using: R_{25}, R_{50}, and R_{90} for the gas and the stars, along with the exponential scale length, $R_{\rm disk}$, from a 2-component (sérsic plus exponential) fit to the surface density. Table 2.1 lists the values for each galaxy. We also compare these scaled

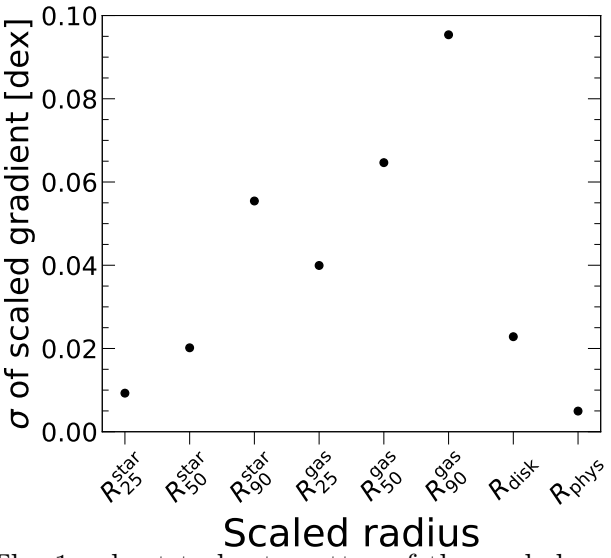


FIGURE A.1. The 1-σ host-to-host scatter of the scaled radial gradients for different scale radii at $z = 0$. We define R_{25}, R_{50}, and R_{90} for the gas and stars in Table 2.1. $R_{\rm disk}$ is the exponential scale length of the stellar disk determined via a 2-component fit to the surface density, and $R_{\rm phys}$ is the physical radial coordinates of the disk, i.e. unscaled coordinates. For each scale radius, we measure the gradient of all galaxies across an equal radial range that corresponds to $4-12$ kpc physical for the galaxy with the median scale length. The 1-σ host-to-host scatter is smallest when measuring the gradients in physical space, which motivates our choice for our analysis in this paper.

gradients to the gradients in physical radii (as measured in Section. 2.4.2. We bin each profile equally in scaled radius, defining the bin width such that the galaxy with the median scale length has a binwidth of 0.25 kpc physical. We measure the radial gradient of each galaxy across an equal radial range for each scale radius. We define this radial range such that we measure the galaxy with the median scale length across a physical range $4 - 12$ kpc. This range corresponds to: $\approx 3.3-9.8R_{25}^{\rm star}$, $\approx 1.4-4.1R_{50}^{\rm star}$, $\approx 0.4-1.1R_{90}^{\rm star}$, $\approx 0.5-1.6R_{25}^{\rm gas}$, $\approx 0.3 - 1.0R_{50}^{\rm gas}$, $\approx 0.2 - 0.7R_{90}^{\rm gas}$, and $\approx 1.0 - 3.1R_{\rm disk}$.

Measuring the gas abundance radial gradient (from $4-12$ kpc) in physical space minimizes the host-to-host scatter, to $\sigma \approx 0.005$ dex. $R_{25}^{\rm star}$ has the next smallest 1-σ scatter with $\sigma \approx 0.009$ dex. The gradients are the least self-similar when scaled by $R_{90}^{\rm gas}$, for which the 1-σ scatter is ≈ 0.1. The self-similarity of the radial profiles in physical space motivates our choice in this paper, because there is no compelling reason to scale the profiles of our galaxies. We emphasize, though, that this may be a result of the small mass range of our suite (halo masses are $M_{\rm 200m} = 1 - 3 \times 10^{12} {\rm M}_{\odot}$, stellar masses are in Table 2.1) and may not be generalizable to galaxies across a wider mass range.

A.2. All gas versus star-forming gas

In this paper, we examine elemental abundances in all gas, as initial conditions for chemical tagging of stars. We choose to measure all gas in part because star-forming gas represents only a small fraction of all gas elements at a given snapshot, leading to significant Poisson noise. In principal, we could attempt to identify photo-ionized (HII) regions near young star particles to compare with gas-phase measurements via nebular emission lines, but doing this correctly requires generating synthetic observations via ray-tracing, which is beyond the scope of our analysis. In future work (Bellardini et al., in prep.) we will compare in detail the spatial variations in abundance of star particles that form out of this gas to the gas itself. Here, we explore the impact of measuring only star-forming gas instead of all gas.

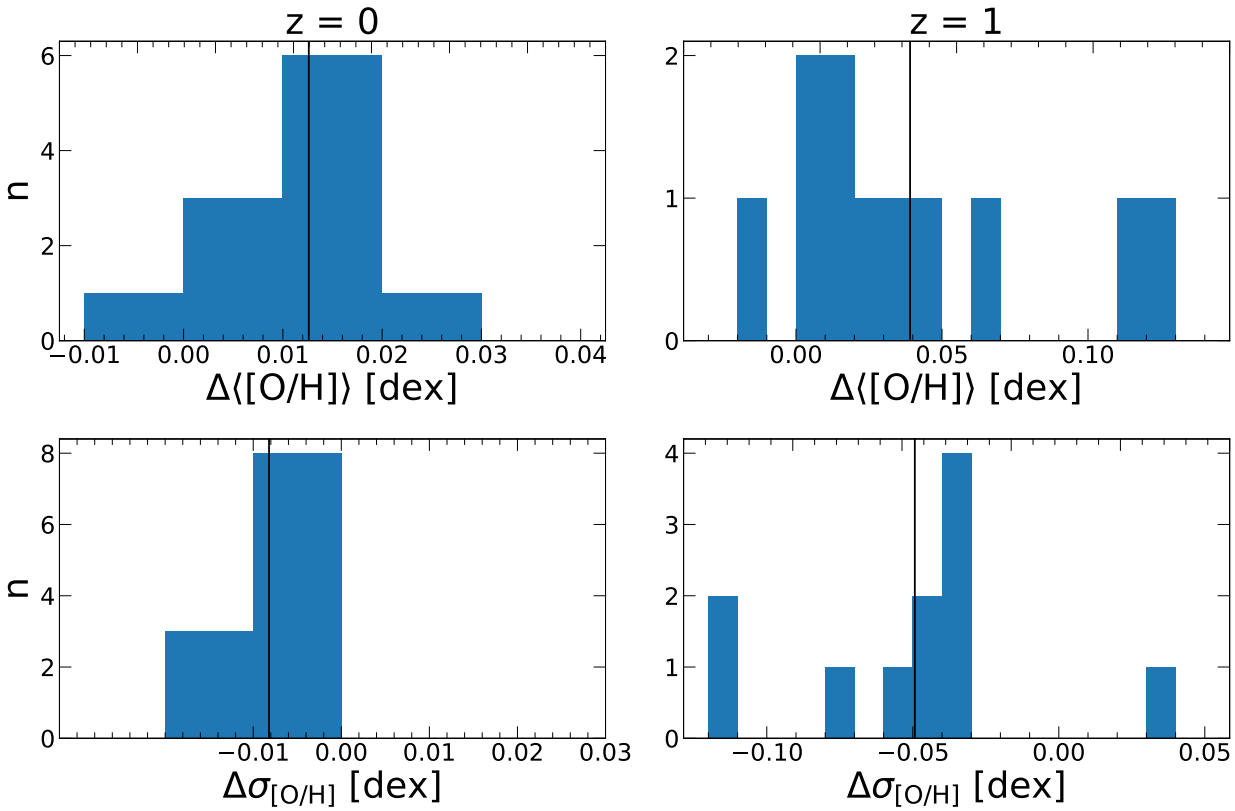


FIGURE A.2. A comparison of [O/H] measured in all gas versus only star-forming gas, at $4 < R < 12\,\mathrm{kpc}$ and $|Z| < 1\,\mathrm{kpc}$. For each host, we measure the mean and standard deviation of its [O/H] (stacking 10 consecutive snapshot across $\approx 200\,\mathrm{Myr}$ to boost the number of star-forming gas elements), and we compute the galaxy-wide difference between star-forming gas and all gas. Top panels show histograms of the difference in the mean [O/H], while bottom panels show histograms of the difference in the standard deviation of [O/H]. Left panels show $z = 0$ and right panels show $z = 1$ ($t_{\mathrm{lookback}} = 7.8\,\mathrm{Gyr}$). The solid vertical lines show the mean of each difference. Star-forming gas is on average more metal rich than all gas by $\approx 0.04\,\mathrm{dex}$ at $z = 1$ and $\approx 0.01\,\mathrm{dex}$ at $z = 0$. Furthermore, star-forming gas is slightly better mixed (with less scatter), with $\sigma_{[\mathrm{O/H}]} \approx 0.05\,\mathrm{dex}$ smaller at $z = 1$ and $\approx 0.008\,\mathrm{dex}$ smaller at $z = 0$.

Fig. A.2 compares measuring [O/H] in star-forming versus all gas at $z = 1$ ($t_{\mathrm{lookback}} = 7.8\,\mathrm{Gyr}$) and $z = 0$. For each galaxy, we select gas elements at $4 < R < 12\,\mathrm{kpc}$ and $|Z| < 1\,\mathrm{kpc}$, and we stack this measurement across 10 snapshots ($\approx 200\,\mathrm{Myr}$), because at any single snapshot there are few star-forming gas elements. For reference, for these

same simulations at $z = 0$, Benincasa et al. (2020) find typical GMC lifetimes, and hence lifetimes of star-forming regions, of $5 - 7\,\mathrm{Myr}$. We first measure the difference in the average abundance between star-forming and all gas for each galaxy. Fig. A.2 (top row) shows a histogram of this offset in the mean [O/H]. A positive value means star-forming gas has a higher [O/H] than all gas for that galaxy. The black vertical line shows the mean value of the histogram. On average, star-forming gas has modestly higher [O/H] than all gas by $\approx 0.04\,\mathrm{dex}$ at $z = 1$ and $\approx 0.01\,\mathrm{dex}$ at $z = 0$. The difference in [O/H] is typically $\lesssim 0.02\,\mathrm{dex}$ for $z = 0$ and always less than $0.03\,\mathrm{dex}$. The discrepancy is larger at higher redshift, the difference is typically $\lesssim 0.04\,\mathrm{dex}$ and always less than $0.13\,\mathrm{dex}$. This is likely because cosmic accretion and star-formation rates are higher at earlier times, leading to less efficient small-scale mixing of gas. Of course, a simple offset in the [O/H] normalization does not alone mean that spatial variations are different.

Fig. A.2 (bottom row) shows the difference in the standard deviation of star-forming versus all gas. Again, the black line shows the mean value. On average, [O/H] for star-forming gas has slightly smaller standard deviation than for all gas. This difference is larger at $z = 1$ than at $z = 0$. However, the difference is typically small, $\lesssim 0.05\,\mathrm{dex}$. This suggests that the azimuthal variations of star-forming gas may be smaller than that of all gas, especially if the scatter is driven primarily by radial variations in abundance. Thus, chemically tagging the birth radii of stars may may be complicated by azimuthal variations for redshifts higher than we show in Sec. 2.4.5, which we will explore further in Bellardini et al. in prep.

We also explore the differences in the radial gradients for star-forming versus all gas (not shown). At $z = 0$ the radial gradients of star-forming and all gas are consistent to within $\pm 0.005\,\mathrm{dex\,kpc^{-1}}$, less than the host-to-host scatter in Fig. 2.5. At $z = 1$, for the galaxies with sufficient star-forming gas to measure a reliable radial gradient, they agree with the gradients for all gas to within $\pm 0.002\,\mathrm{dex\,kpc^{-1}}$. We also compare compare the radial profiles

of [O/H] for newly formed stars (in age bins of 200 Myr) with that of all gas at the same snapshots. For $z \lesssim 0.5$, these profiles overlap to within uncertainty. For $z \gtrsim 0.5$, the stellar and gas abundance profiles start to diverge, such that the young stars tend to have higher [O/H] and flatter gradients than all gas, which we will explore further in Bellardini et al., in prep.

In summary, while analyzing all gas is a reasonable, if not perfect, proxy for star-forming gas in our simulations. Given the short lifetimes of star-forming gas clouds (Benincasa et al., 2020) and the strict conditions for particles to be star-forming (Hopkins et al., 2018), only a small fraction of gas elements are star-forming at a given snapshot ($\lesssim 1\%$ of gas particles at the redshifts we observe), so analyzing all gas greatly reduces the statistical uncertainty. In the future we will study the abundances of star particles that form from this gas, to compare with these results in detail.

A.3. Impact of diffusion coefficient

Our FIRE-2 simulations model the sub-grid diffusion/mixing of metals in gas via unresolved turbulent eddies (Su et al., 2017; Escala et al., 2018; Hopkins et al., 2018):

$$\frac{\partial Z_i}{\partial t} + \nabla \cdot (D \nabla Z_i) = 0 \tag{A.1}$$

where Z_i is the mass fraction of a metal in gas element i, and D is the diffusion coefficient. While there is some uncertainty in the exact value to choose for this coefficient, our fiducial value is physically motivated based on tests of the metal diffusion implementation in FIRE-2 on idealized, converged turbulent box simulations by Colbrook et al. (2017) and other more extensive studies by Rennehan et al. (2019). Here, we compare our key results using our fiducial diffusion coefficient D in m12i against a re-simulations of m12i with all identical physics/parameters, except one has a diffusion coefficient that is 10 times higher (that is, faster mixing) and the other includes no subgrid mixing.

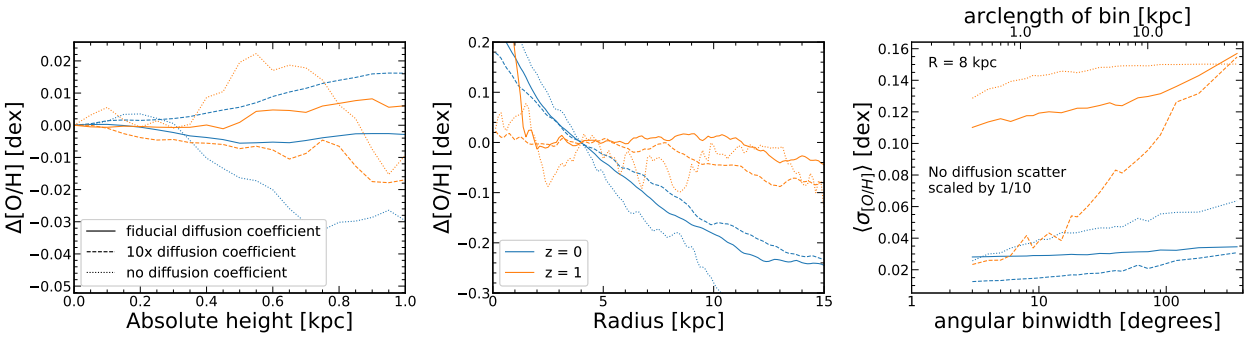


FIGURE A.3. Vertical (left), radial (middle), and azimuthal variations (right) in [O/H] between our fiducial simulation of m12i, a version with no subgrid metal diffusion, and a re-simulation increasing the diffusion coefficient by 10 times. The vertical profiles show no clear systematic variations at a level important for our analysis. We normalize the radial profiles to 4 kpc (the approximate edge of the bulge) for clarity in comparison. The radial gradients (measured from $4-12$ kpc for $z=0$, $2-8$ kpc for $z=1$) vary by no more than $\approx 0.005\,\mathrm{dex\,kpc^{-1}}$ between our fiducial simulation and the simulation with 10 times higher metal diffusion, while the simulation with no metal diffusion has a $\approx 0.13\,\mathrm{dex\,kpc^{-1}}$ steeper gradient at $z=0$. In the right panel, we scaled down the azimuthal scatter in the simulation with no metal diffusion by a factor of 10, for comparison. Thus, neglecting metal diffusion/mixing leads to $10\times$ higher azimuthal scatter, and moreover, this scatter does not depend much on azimuthal scale. The enhanced metal diffusion re-simulation shows smaller azimuthal scatter at small azimuthal scales, given the enhanced mixing rate on these small scales. However, that simulation shows similar scatter at large azimuthal scales, indicating that disk-wide azimuthal scatter is not sensitive to the detailed choice of diffusion coefficient.

Fig. A.3 compares the vertical, radial, and azimuthal variations for m12i. The left panel shows the vertical gradient in gas [O/H], similar to Fig. 2.6. At $z=0$, we find no differences within 200 pc and at most ~ 0.015 dex difference at 1 kpc. The differences are stronger at $z=1$ for $10\times$ higher diffusion and stronger at $z=0$ for the simulation with no diffusion, though again, not at a significant level to change our interpretations, especially within 200 pc.

Fig. A.3 (center) shows the radial profile in gas [O/H], normalized to the abundance at $R=4$ kpc (given a strong upturn at smaller R). The radial gradients, measured over our fiducial radial ranges, vary by $\lesssim 0.005\,\mathrm{dex\,kpc^{-1}}$ between the $10\times$ diffusion simulation and the fiducial simulation. The gradients vary by less than 0.014 dex between the fiducial

simulation and the one with no metal diffusion at $z = 0$. The simulation with no subgrid diffusion has a steeper gradient at $z = 0$, potentially because the metals are less efficient at spreading from a given radius, in the absence of subgrid diffusion, once the disk has become rotationally dominated and radial turbulence is no longer efficient at moving the gas particles. We thus conclude that the radial gradients are reasonably robust to choices of the strength of the diffusion coefficient, however, in the unphysical case of no subgrid diffusion, the gradient can be (unphysically) steeper.

Fig. A.3 (right) compares the azimuthal variations versus angular bin width, at $R = 8\,\mathrm{kpc}$. The simulation with no subgrid diffusion has $10\times$ higher azimuthal scatter, so we scale down its values in Fig. A.3 by $10\times$ for visual comparison. Using no subgrid diffusion leads to scatter that is largely independent of azimuthal scale at high z, but that increases with azimuthal scale at low z. Without subgrid diffusion, a small number of gas elements can absorb most of the metals, while a significant number of (neighboring) elements can remain nearly un-enriched. This is a patently unphysical scenario, and it yields azimuthal scatter that disagrees with observations by an order of magnitude. At both redshifts, using a higher diffusion coefficient leads to smaller azimuthal variations at small scales, because diffusion smooths variations between nearby gas elements on scales approaching the resolution (Escala et al., 2018). However, the azimuthal variations are nearly unchanged on large azimuthal scales. Therefore, our results on small azimuthal scales are likely sensitive to the exact choice of diffusion coefficient, but the large-scale azimuthal variations are robust. An important caveat to this comparison is that it is only one simulated galaxy, and individual simulations with the same initial conditions and physics can show non-trivial stochastic variations from random number generators, floating-point roundoff, and chaotic behavior (e.g. Keller et al., 2019). Indeed, we find minor fluctuations between these simulations, for example, in the exact timing of mergers, which can affect all panels in Fig. A.3. We consider it likely that

the differences in azimuthal variations on small scales are robust, but any other differences in Fig. A.3 are potentially stochastic.

APPENDIX B

Chapter 3 Appendices

B.1. Shapes of abundance profiles

We investigate the cause of the break in the radial abundance profiles (see Fig. 3.3) by exploring the profiles of surface density. We fit a two-component linear profile to the log of the young stellar to gas surface ratio in Fig. 3.4 (bottom) and plot the transition radii in Fig. B.1 in green. In principal, the ratio of surface density of stars to gas should be approximately proportional to the abundance profile, in the limit of local metal deposition.

The blue points in Fig. B.1 show the transition from the steep inner radial abundance profile to the flatter outer abundance profile for each galaxy as a function of $M_{90}^{\rm star}$ (see Table 3.1 for masses), and the horizontal line shows the average of 5.8 kpc across all 11 galaxies. The orange points show the radius at which the overall stellar surface density transitions from being dominated by a sérsic profile to being dominated by an exponential profile, which we presented in B21, and which reflects the transition to a bulge-like component in each galaxy. We found this by simultaneously fitting an exponential plus sérseic profile to the overall stellar surface density at $z = 0$, where we fixed the sérsic index at $n = 1.3$.

Fig. B.1 shows no clear stellar mass dependence to these transition radii. The transition radius in [Fe/H] for young stars is always larger than the transition radius in the surface density, typically by ~ 2.6 kpc. Additionally, the transition radius of the ratio of stellar to gas surface density is, on average, larger than transition radius of the abundance profile by ~ 1.3 kpc.

We thus conclude that the transition radius in the abundance profile for young stars in our simulation is not related to the onset of a bulge-like component. Instead, it coincides

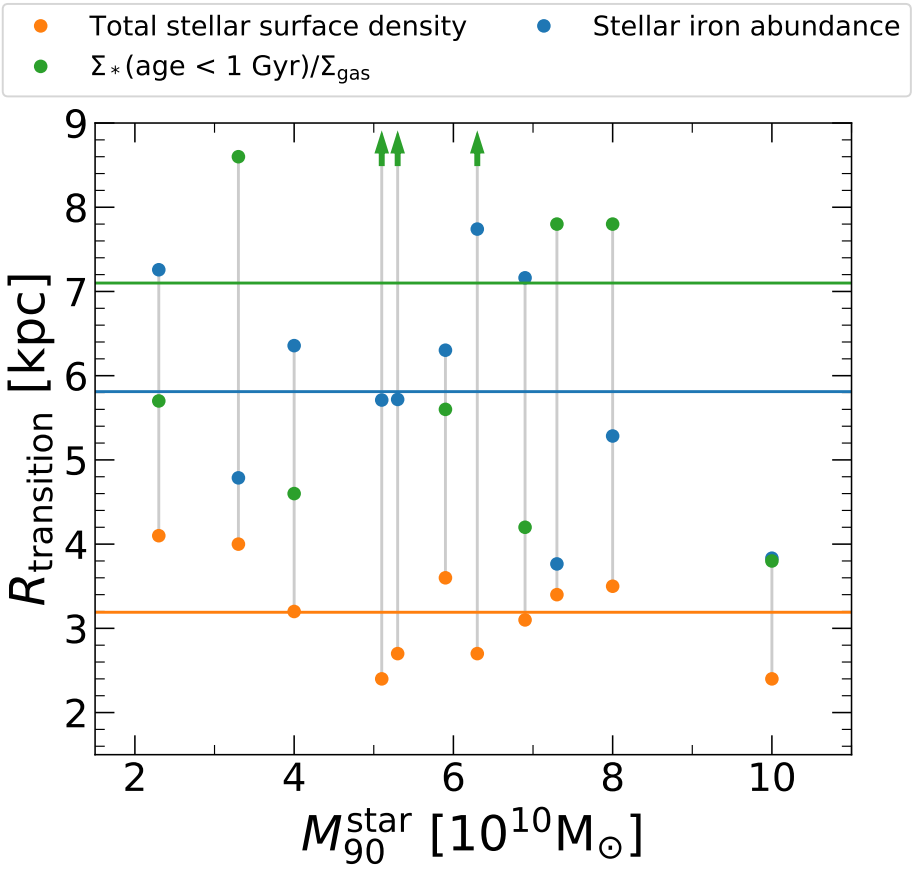


FIGURE B.1. Various transition (break) radii ($R_{\rm transition}$) for each galaxy as a function of stellar mass. The orange points show $R_{\rm transition}$ of the stellar surface density profile, caused by the steeper bulge-like region, as defined in B21. The blue points show $R_{\rm transition}$ from a 2-component linear fit to the [Fe/H] profiles, as in Fig. 3.3. The green points show $R_{\rm transition}$ from a 2-component linear fit (measured from $R = 3 - 12\,$kpc) to ratio of the stellar surface density to gas surface density for stars younger than 1 Gyr old. The horizontal lines show the mean of each $R_{\rm transition}$. $R_{\rm transition}$ as defined by the stellar iron abundance is always larger than $R_{\rm transition}$ as defined by the surface density of all stars. We expect a correlation between $R_{\rm transition}$ for the stellar surface density and $R_{\rm transition}$ for the stellar [Fe/H] because the [Fe/H] abundance of the youngest stars is a result of the supernovae from previous stellar generations.

better with the transition radius in the stellar to gas ratio of surface densities, and its shape (break) is at least partially set by the shape of that ratio. However, we do not find perfect agreement. More work is needed to understand the full shape of the abundance gradient in the context of complex metal injection, mixing, outflows, and stellar redistribution.

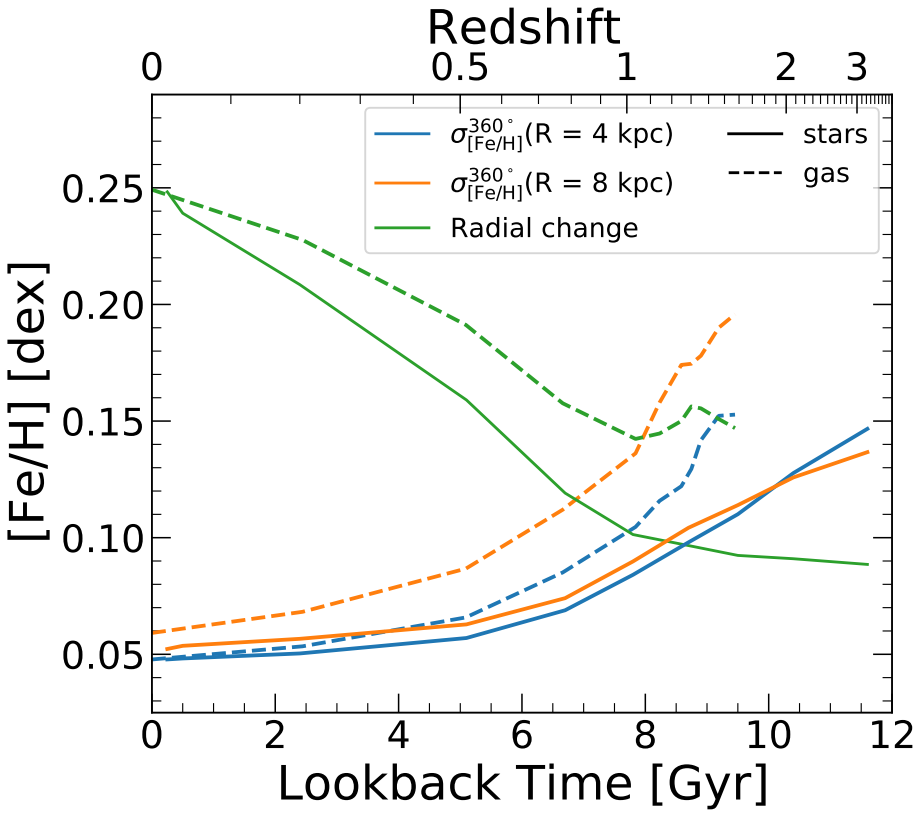


FIGURE B.2. Similar to Fig. 3.11, comparing the results for elemental abundance variations in gas versus young (age $< 500\,\mathrm{Myr}$) stars. We generate the profiles in the same way as in Fig. 9 of B21 and we smooth all profiles with a Gaussian filter. In general, both the azimuthal scatter and the radial change are smaller for newly formed stars than for all gas. Also, the transition time (see Section 3.4.7) is generally earlier for young stars than for all gas.

B.2. Stars at formation versus gas

We compare the 360° azimuthal scatter and the strength of the radial gradient of newly formed stars (age $< 500\,\mathrm{Myr}$) to that of all gas, as we presented in B21. For consistency with B21, instead of examining the radial change in [Fe/H] from $R = 0\,\mathrm{kpc}$ to R^*_{90}, we measure the radial gradient as a linear profile from $R = 4\,\mathrm{kpc}$ out to $R = 12\,\mathrm{kpc}$ and multiply by $8\,\mathrm{kpc}$ to define a radial change in [Fe/H].

Fig. B.2 shows the mean radial change (green) and the 360° scatter at $R = 4\,\mathrm{kpc}$ (blue) and $8\,\mathrm{kpc}$ (orange) for young stars (solid) and gas (dashed). We find excellent agreement at

$z \sim 0$, which we confirm by comparing the radial profiles of young stars and gas directly. However, young stars show systematically lower variations, both radial and azimuthal, at all lookback times. The largest discrepancy in the radial gradient is $\approx 0.009\,\mathrm{dex\,kpc^{-1}} \approx 9.5\,\mathrm{Gyr}$ ago. We believe this is because the stellar gradients are ill defined in this radial range for young stars at this lookback time. As Fig. 3.1 shows, R^*_{90} is only $\approx 5.5\,\mathrm{kpc}$ on average. The greater agreement between azimuthal scatter for gas and stars at 4 kpc compared to 8 kpc also supports this. As mentioned in Section 3.4.6, we think that clustered star formation drives smaller azimuthal scatter in young stars than in all gas. B21 showed that star-forming gas in general has a smaller disk-wide scatter in abundance (median of $\sim 0.05\,\mathrm{dex}$ at $z = 1$). This agrees with the typical difference in azimuthal scatter at large lookback times in Fig. B.2.

B.3. Measurable homogeneity

Following B21, we define $\Delta R_{\mathrm{equality}}$, the ratio of the (radially averaged) 360° azimuthal scatter to the overall radial gradient:

$$\Delta R_{\mathrm{equality}} = \frac{\sigma_{[\mathrm{X/H}]}}{\Delta[\mathrm{X/H}]_{R^*_{90}}/R^*_{90}} \tag{B.1}$$

$\Delta R_{\mathrm{equality}}$ in effect defines the minimum radial scale over which radial variations dominate over azimuthal variations in abundance. This sets a maximum precision that chemical tagging neglecting azimuthal scatter can place on the birth radius of a star using a given abundance measurement of the star.

Fig. B.3 shows the median $\Delta R_{\mathrm{equality}}$ as a function of lookback time. The blue line shows the median ratio for [Fe/H], the orange line shows the median ratio for [Mg/H], and the black line shows the average R^*_{90} of newly formed stars, from Section 3.4.1. The shaded region shows the $1-\sigma$ scatter for [Fe/H]. $\Delta R_{\mathrm{equality}}$ in general decreases with decreasing lookback time ($\approx 1.2\,\mathrm{kpc}$ at present day and $\approx 11.2\,\mathrm{kpc}$ 12 Gyr ago), meaning that the most precision can be placed on the birth radii of stars formed within the last $\sim 2\,\mathrm{Gyr}$.

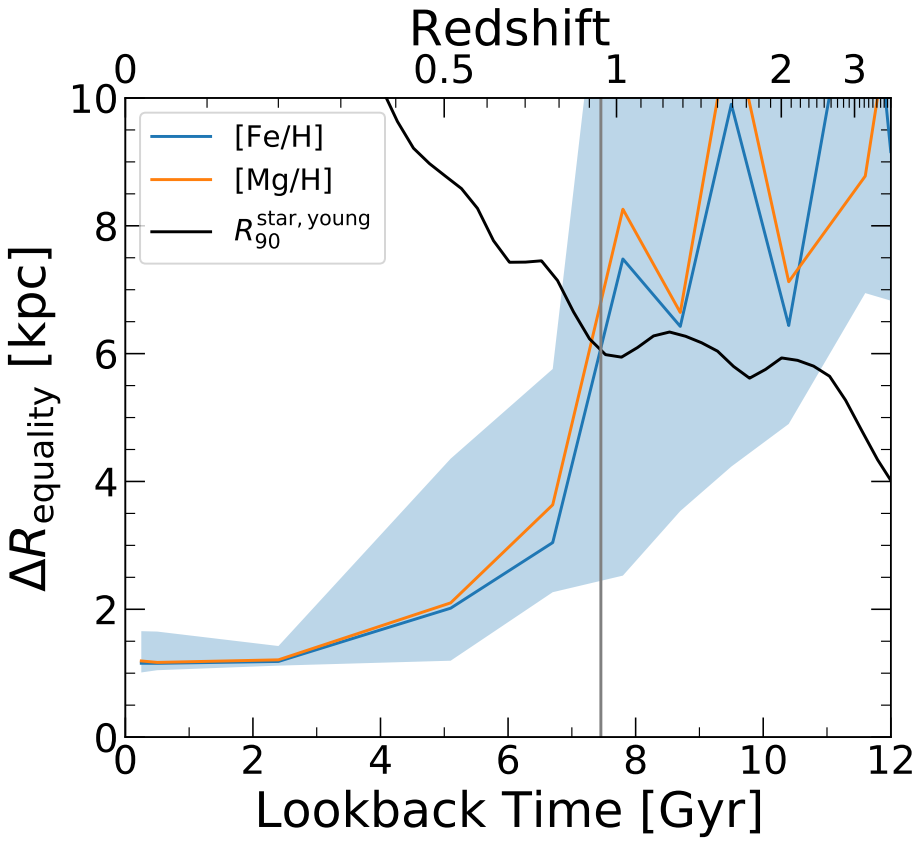


FIGURE B.3. The lines show the median $\Delta R_{\text{equality}}$ as a function of lookback time for [Fe/H] (blue) and [Mg/H] (orange) and the shaded region shows the 68th percentile for [Fe/H]. For comparison, the black line shows the average R_{90}^{*} of newly formed stars at each lookback time. We define $\Delta R_{\text{equality}}$ as the ratio of the 360° azimuthal scatter to the radial gradient. This ratio gives a radial scale over which azimuthal variations dominate over radial variations for [Fe/H] (blue) and [Mg/H] (orange). $\Delta R_{\text{equality}}$ effectively sets the precision to which birth radii of stars can be measured if azimuthal scatter is neglected. For stars that formed prior to $\approx 7.5\,\text{Gyr}$ ago (right of the gray line), the azimuthal scatter dominated over radial variations across the entire galaxy. However, for stars that formed within the past $\sim 6\,\text{Gyr}$, azimuthal scatter only dominates on scales $\lesssim 3\,\text{kpc}$, less than half of R_{90}^{*}.

We can, in principal, infer the birth radii of all stars born at $t_{\text{lb}} \lesssim 6\,\text{Gyr}$ to within 3 kpc. For stars born at $t_{\text{lb}} \lesssim 500\,\text{Myr}$, the uncertainty is $\approx 1.1\,\text{kpc}$ using just measured [Fe/H] or [Mg/H]. For stars born at $t_{\text{lb}} \gtrsim 7.5\,\text{Gyr}$ (right of the gray line), $\Delta R_{\text{equality}}$ was larger than the size of the galaxy. This is effectively another way to define the transition age in Fig. 3.11.

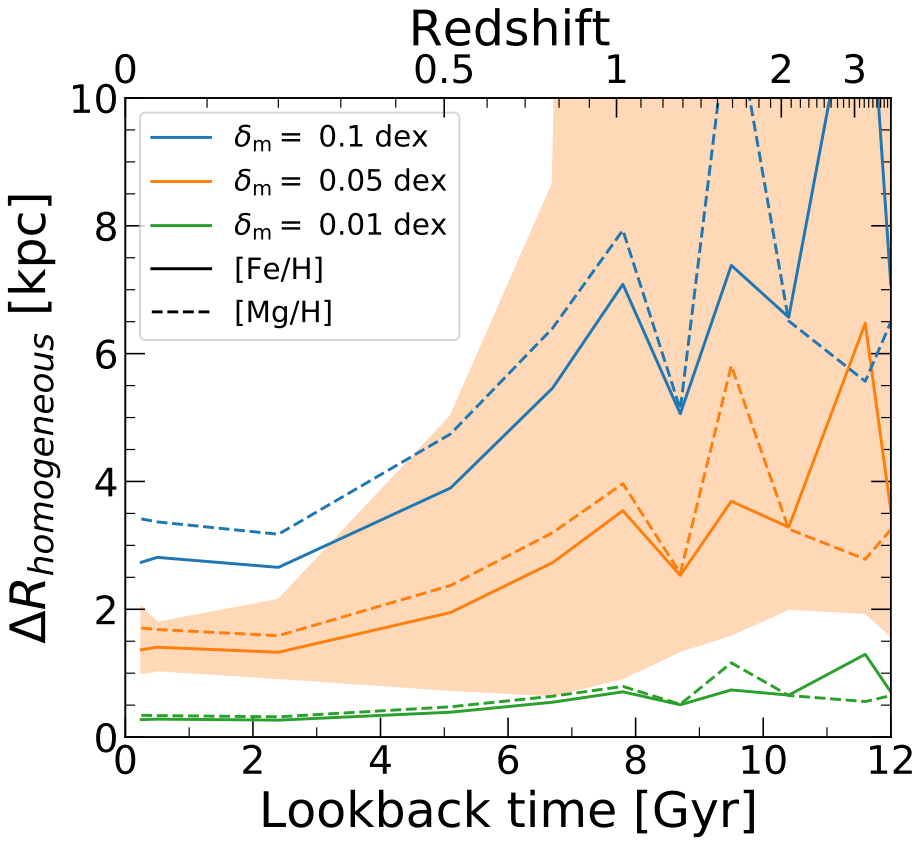


FIGURE B.4. $\Delta R_{\rm homogenous}$ as a function of lookback time. $\Delta R_{\rm homogenous}$ is the ratio of assumed measurement uncertainties (taken as representative of typical observational uncertainties) to the fiducial radial gradients in our 11 simulated galaxies. The solid (dashed) line shows the median value for the simulations for [Fe/H] ([Mg/H]). The orange shaded region shows the full distribution for a fiducial scatter of 0.05 dex. This ratio predicts the precision to which the radial birth location of a star is definable, for a given measurement precision, assuming the radial abundance gradient is primarily responsible for setting the abundance of a star. In the simplified case of no azimuthal scatter, measurement uncertainty sets the precision of inferred stellar birth radii. Which, for our fiducial uncertainty of $\delta_{\rm m} = 0.05$ dex, is $\Delta R_{\rm homogeneous} \lesssim 2.7$ kpc for $t_{\rm lb} \lesssim 8.7$ Gyr.

Of course, one may be able to improve on this precision using multiple abundances at once, which we will explore in future work.

Additionally, we examine the effect of observational measurement uncertainty on the precision with which chemical tagging can indicate stellar birth radius. We present results for several measurement uncertainties $\delta_{\rm m} = 0.1$, 0.05, and 0.01 dex, representative of low-,

medium-, and high-resolution spectroscopic surveys, for example, GALAH (Buder et al., 2021). We define $\Delta R_{\rm homogeneous}$ as the ratio of the measurement uncertainty to the radial gradient:

$$\Delta R_{\rm homogeneous} = \frac{\delta_{\rm m}}{\Delta[{\rm X/H}]_{R^*_{90}} / R^*_{90}} \tag{B.2}$$

This ratio defines the radial scale over which the stellar disk is measurably homogeneous, assuming only a radial abundance gradient. Thus, it defines the measurement-limited precision on the birth radius of a star born in a disk dominated by a radial abundance gradient.

Fig. B.4 shows $\Delta R_{\rm homogeneous}$ as a function of lookback time. The solid lines show the median for [Fe/H], and the dashed lines show the median for [Mg/H]. We show the full distribution for our fiducial uncertainty of $\sigma_{\rm m} = 0.05\,$dex. For all lookback times, medium-resolution surveys give $\Delta R_{\rm homogeneous} \lesssim 4.4\,$kpc, with all stars formed at $t_{\rm lb} \lesssim 8.7\,$Gyr having $\Delta R_{\rm homogeneous} \lesssim 2.7\,$kpc.

Fig. B.4 and Fig. B.3 show that $\Delta R_{\rm equality} \lesssim \Delta R_{\rm homogeneous}$ for all lookback times $\lesssim 6.7\,$Gyr for our fiducial measurement uncertainty of $0.05\,$dex. Thus, measurement uncertainty is the limiting factor in setting the precision of birth radii for stars born less than $\sim 5\,$Gyr ago. However, for stars born at $t_{\rm lb} \gtrsim 6.7\,$Gyr, $\Delta R_{\rm equality}$ was $2-3\times$ larger, so the azimuthal scatter then was more important in setting the precision of stellar birth location.

APPENDIX C

Chapter 4 Appendices

C.1. Effect of Radial Selection

The selection criterion we use in this paper, when plotting radially averaged trends (right panel of Fig. 4.2 and Fig. 4.3), is $2 \leq R_{\rm phys}^{\rm now} \leq 12\,{\rm kpc}$. The upper bound is well justified by the scale radii of the disks (see Bellardini et al., 2021, Table. 1). However, the lower bound is less well defined. We select a lower bound in this paper based on the radii at which stars

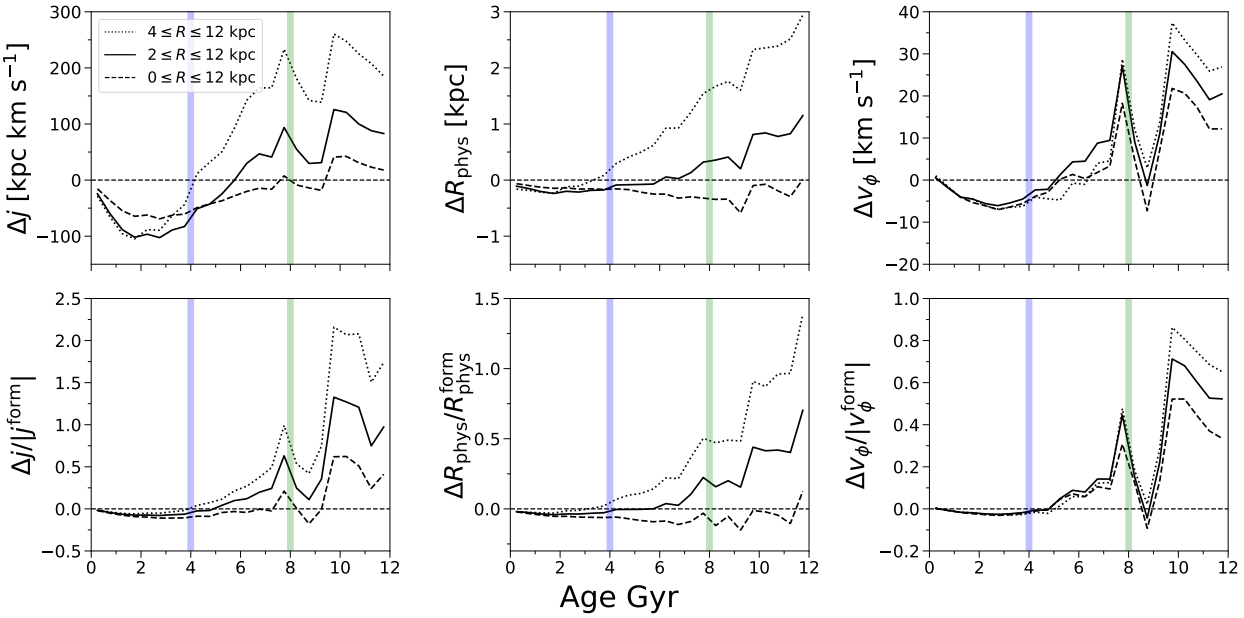


FIGURE C.1. The same as Figure 4.3, except the different lines show trends selecting stars across different radii; the solid lines show our fiducial selection. Because of the declining density profile of stars in the galaxy, excluding smaller radii (including the bulge region) significantly changes the population of stars included. In general, the absolute and fractional change of angular momentum, radius, and azimuthal velocity for older stars is significantly larger when excluding stars in the inner galaxy, while this radial selection has a weaker effect for younger stars.

generally form on circular orbits, but here we test the robustness of our results to that choice by redoing our analysis including stars from $0-12\,\mathrm{kpc}$ and from $4-12\,\mathrm{kpc}$.

We do not show the results of this analysis for Fig. 4.2 (right) as we find essentially no change to the figure when varying our selection region. Selecting all stars in the disk skews the trend line down by an average of $\approx 5\%$ and selecting stars from $4-12\,\mathrm{kpc}$ skews the trend line up by $\approx 3\%$. The shape of the trend is unchanged. The variation is most noticeable in stars older than $\approx 8\,\mathrm{Gyr}$ as there is a strong radial dependence to the circularity of the oldest stars in the disk.

In Fig. C.1 we recreate Fig. 4.3 using our different selection regions. The solid black line is our fiducial selection region, the dashed black line shows the trends for stars with $0 \leq R_{\mathrm{phys}}^{\mathrm{now}} \leq 12\,\mathrm{kpc}$, and the dotted black lines shows the trends for stars with $4 \leq R_{\mathrm{phys}}^{\mathrm{now}} \leq 12\,\mathrm{kpc}$.

The most important takeaway from Fig. C.1 is the qualitative similarity of all of the results regardless of our radial selection. The only qualitative difference we see is in the change in radius and the fractional change in radius for stars from $0-12\,\mathrm{kpc}$. There is little to no radial redistribution when including all stars, whereas the radial redistribution increases with age when excising the inner disk. The negative density gradient of galaxies results in properties of the inner galaxy biasing results presented as a disk-wide average. This, coupled with a predominantly older stellar population in the galactic center and our scalefactor dependent in-situ cut leads to a net decrease in the average radial redistribution of the oldest stars when including stars presently in the innermost galaxy. Our selection criterion guarantees that the oldest stars presently near the center of the galaxy necessarily formed there.

C.2. Binning stars by physical radius

Fig. C.2, is the same as Fig. 4.7 except the stars are binned by $R_{\mathrm{phys}}^{\mathrm{now}}$ (width of $0.5\,\mathrm{kpc}$). Thus, the stellar populations are consistent for each metric of radius.

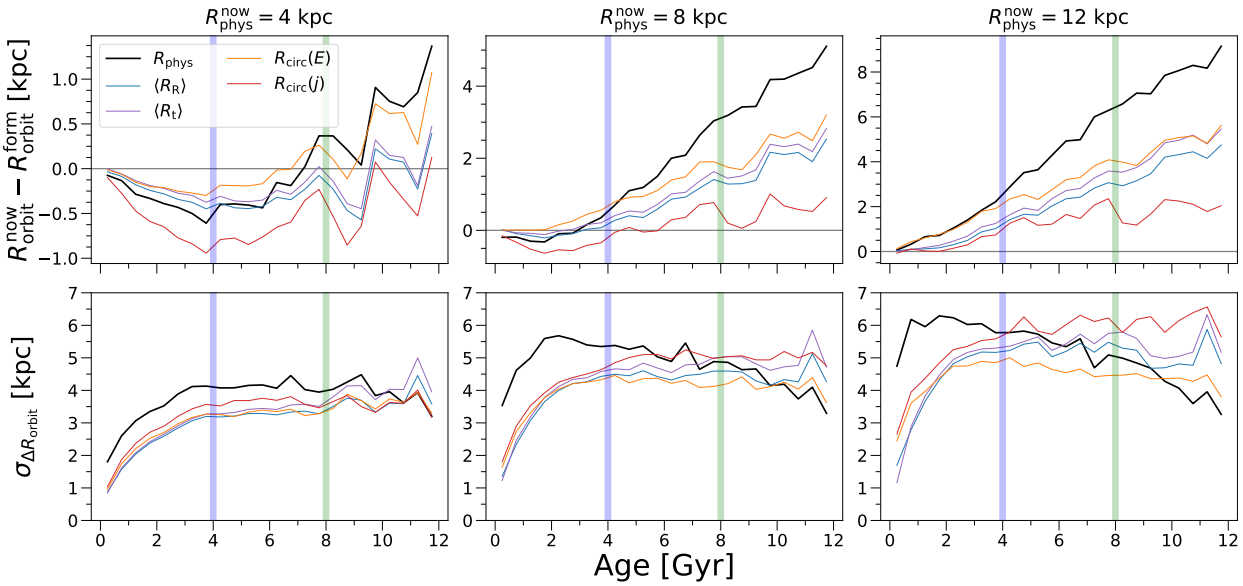


FIGURE C.2. Similar to Figure 4.7, the median change in radius from birth to today versus stellar age, using different metrics of orbital radius to measure radial redistribution, but in all cases electing stars via $R_{\rm phys}$ at $z = 0$. When selecting stars today using a different radius metric than for measuring radial redistribution, we find significantly less agreement between the radius metrics. This motivates the importance of both selecting stars today and measuring radial redistribution self-consistently.

The trends in Fig. C.2 (top) are qualitatively consistent with Fig. 4.7 (top). However, the results are quantitatively inconsistent. $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ tends to decrease for all ages and radial metrics when selecting stars based on their physical radii. This makes sense given the results of Fig. 4.1 as the physical radii are systematically larger than $R_{\rm orbit}$ for nearly all radii.

For all radii and for all ages $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$ tends to be most positive for $R_{\rm phys}$ and most negative for $R_{\rm circ}(j)$. The trends with other radial metrics are all in good agreement with one another, and fall between these two extrema.

The scatter in $R^{\rm now}_{\rm orbit} - R^{\rm form}_{\rm orbit}$, for stars binned by $R^{\rm now}_{\rm phys}$ is qualitatively similar to that of stars binned by $R^{\rm now}_{\rm orbit}$. However, the scatter is generally larger for the stars binned by $R_{\rm phys}$.

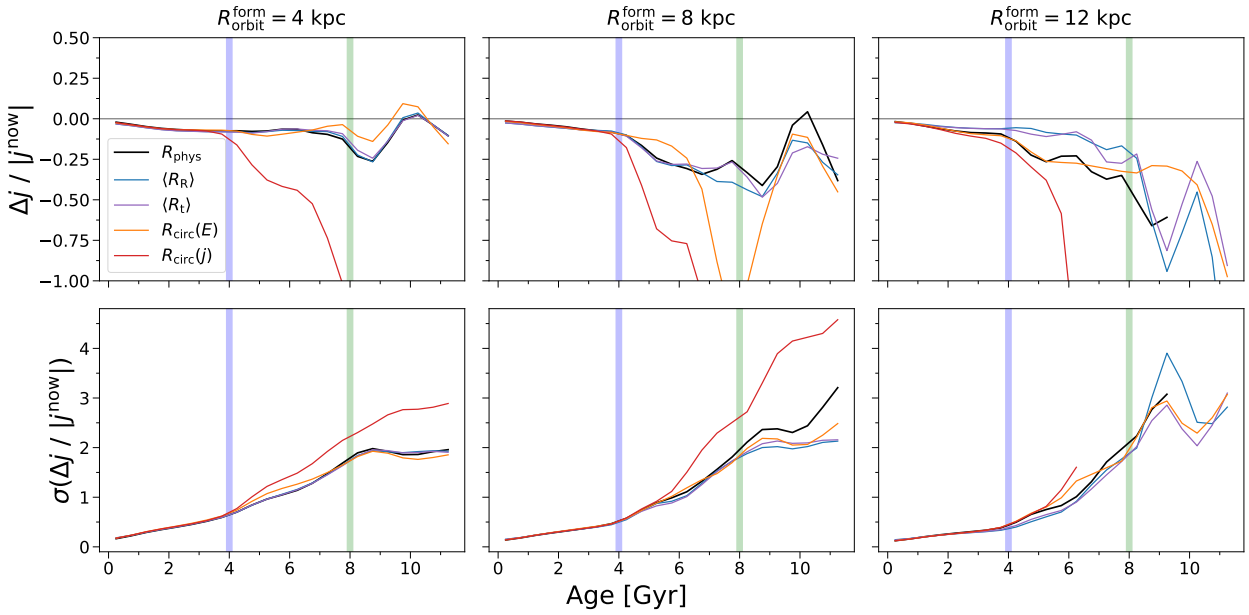


FIGURE C.3. Similar to Figure 4.4, but selecting stars based on formation radius. In general, stars that formed at a given radius have lost orbital angular momentum. The fractional change is approximately constant with age for stars that formed in the inner galaxy, but it grows in magnitude with age for stars that formed at larger radii.

The difference is typically less than $\approx 1\,\mathrm{kpc}$. However, the difference is larger for $R_{\mathrm{circ}}(j)$ at $R_{\mathrm{phys}}^{\mathrm{now}} = 12\,\mathrm{kpc}$, reaching as large as $\approx 5\,\mathrm{kpc}$ for the oldest stars.

C.3. Selecting by Formation Radius

As a test of the similarity between our different metrics of orbital radii and the physical radii of stars, we compare $R_{\mathrm{orbit}}^{\mathrm{form}}$ to $R_{\mathrm{phys}}^{\mathrm{form}}$ for different age stars as a function of $R_{\mathrm{phys}}^{\mathrm{form}}$ (not shown). The results are qualitatively similar to the results presented in Fig. 4.1.

Stars currently at larger radii show less similarity between their orbital radii and their physical radius, with their orbital radii being systematically smaller than their physical radii. This is also true at formation. Stars which formed at larger radii have larger differences between their orbital radii and their physical radius. Additionally, the physical radius of stars is systematically larger than their orbital radii at formation. The youngest stars show

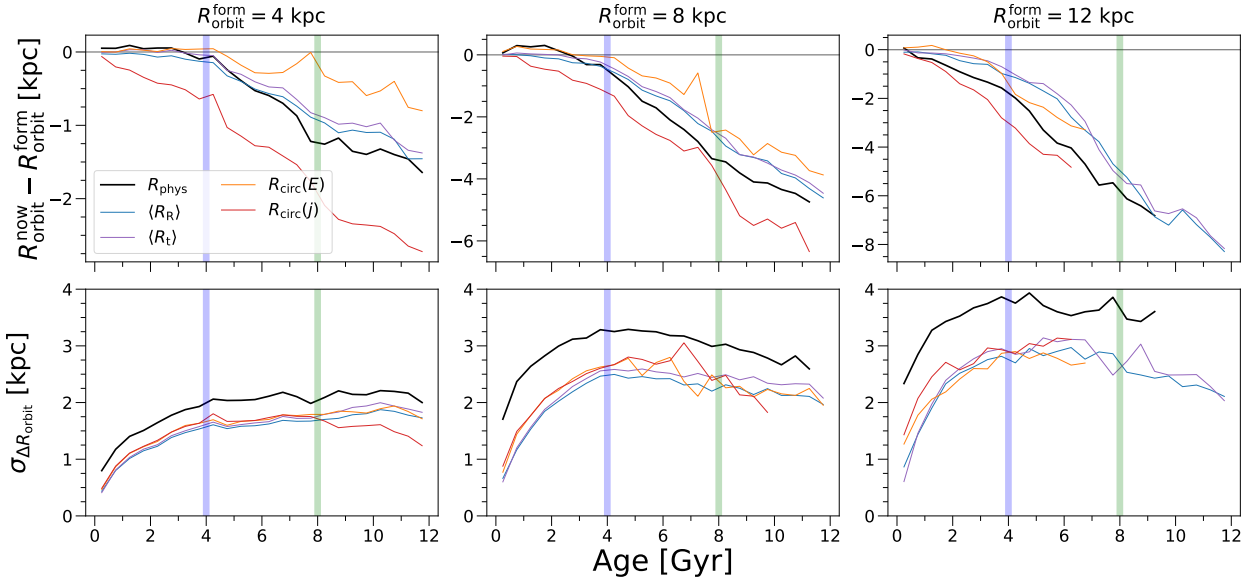


FIGURE C.4. Same as Figure 4.7, but selecting stars in bins of orbital radius at formation $R^{\rm form}$. Now, the median change in orbital radius (top) is almost always negative, with stronger inward redistribution for older stars. But as in Figure 4.7, the scatter (bottom) increases with age to $2-3$ Gyr then becomes nearly independent of age for older stars.

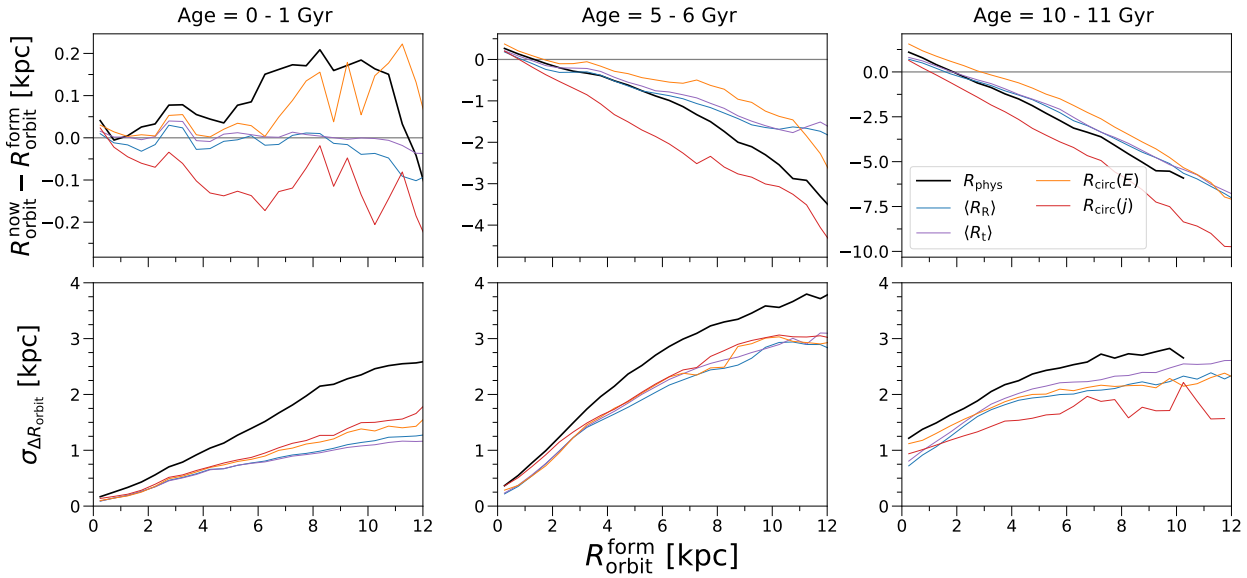


FIGURE C.5. Same as Figure 4.6, but selecting stars in bins of orbital radius at formation $R^{\rm form}$. Now, the median change in orbital radius (top) is generally always negative, with stronger inward redistribution for stars that formed at larger radii. But as in Figure 4.6, the scatter (bottom) increases only mildly with formation radius.

the most similarity between their orbital radii and their physical radii. This reinforces the idea that as disks settle stars form on increasingly circular orbit.

Similar to Fig. 4.4, in Fig. C.3 we measure the fractional change in angular momentum of stars as a function of age, but for stars selected by their formation radii. The different colored lines show the impact of binning by different radial metrics.

Fig. C.3 shows that regardless of formation radius, the median change in angular momentum of stars is negative. Regardless of formation radius, the fractional change in angular momentum is systematically smaller for younger stars. This agrees with the results of Fig. 4.4, and is consistent with disk settling and stars forming on preferentially more circular orbits.

The dependence of the fractional change in angular momentum on age depends on the formation radius. There is little to no dependence on age for stars born in the inner disk; the dependence on age is stronger for stars born at larger radii. The oldest stars born at the largest radii experience the largest fractional changes to their angular momentum.

Regardless of the metric of radius used to select stars, the qualitative changes to the fractional angular momentum with age are consistent. However, the quantitative trends are inconsistent with stars selected via $R_{\rm circ}(j)$ systematically showing larger fractional changes to their angular momentum.

Fig. C.4 is the same as Fig. 4.7, except the stars are selected via their orbital formation radii, as opposed to their orbital radii at $z = 0$. The different colored lines are the simulation averaged median change in radius (top) and the simulation averaged scatter in radial change (bottom) as a function of stellar age for stars selected via different metrics of radius.

Regardless of the chosen metric of radius, the qualitative trends are similar. The absolute radial change of stars tends to increase with increasing formation radius and with increasing stellar age. However, where stars selected by radii at $z = 0$ tend to show positive changes to their radii, stars selected by radii at formation tend to show negative changes to their

radii. The youngest stars show near zero or slightly positive changes to their radii, but for all stars older than a few Gyr, the median change in radius is systematically negative. Thus stars typically redistribute inward, so the positive redistribution of stars seen when selecting them by $R_{\rm orbit}^{\rm now}$ in Fig. 4.7 stems from the density profile of the galaxies.

Stars selected by $R_{\rm circ}(j)$ show the largest systematic change in radius, which is consistent with Fig. 4.7. Stars selected via the other radial metrics generally show qualitative agreement with one another.

Fig. C.4 (bottom) shows the age dependence of the scatter in the radial change of stars. The scatter generally increases with age for stars younger than $\approx 4\,{\rm Gyr}$, and is roughly constant with age for stars older than $\approx 4\,{\rm Gyr}$. This may be tied to the different epochs of disk formation, stars born during the late-disk era generally show less radial redistribution and less scatter to their radial redistribution. Stars born prior to the late disk era are subject to different mechanisms of radial redistribution and thus experience more redistribution and more scatter in their redistribution.

The scatter also shows a radial dependence. For a given age stellar population, the scatter tends to be largest for stars born at the largest radii.

Fig. C.5 is similar to Fig. 4.6, except $R_{\rm orbit}^{\rm now} - R_{\rm orbit}^{\rm form}$ is plotted as a function of $R^{\rm form}$. Fig. C.5 (top) shows that young stars (ages $< 1\,{\rm Gyr}$) typically experience little to no radial redistribution, regardless of their birth radii. However, intermediate and old age stars typically experienced inward redistribution. The magnitude of radial redistribution for older stellar populations increases with increasing formation radius. Additionally, the magnitude of radial redistribution increases with increasing stellar age, the oldest stars have redistributed the most.

The negative radial redistribution typically seen in older stellar populations is not consistent across all radii. Stars born in the inner $1 - 2\,{\rm kpc}$ of the galaxy show positive radial

redistribution on average. This makes sense given they have a vast range of larger radii to which they can be scattered to and essentially no smaller radii to which they can redistribute.

Fig. C.5 (top) also shows that there is qualitative consistency between the trends for the different radial metrics. While not shown here, we also test the trends as a function of $R_{\rm phys}^{\rm form}$. The trends are qualitatively consistent with those present here, however there is much less quantitative agreement between the different radial metrics.

While we do not present results in the same way, the results of Fig. C.5 show qualitative agreement with Okalidis et al. (2022) (Fig. 6). They find that, for a given birth radius, the peak present-day radius of stars decreases with increasing stellar age, i.e. older stars redistribute inward more than younger stars. Fig. C.5 shows that for a given formation radius, older stars experienced more negative redistribution than younger stars.

Fig. C.5 (bottom) shows the 68th percentile scatter in the radial redistribution of the top panel, as a function of orbital formation radius. There is generally quantitative consistency between the different radial metrics, however the scatter in $R_{\rm phys}$ is systematically larger than the scatter in the other radial metrics.

For all metrics of radius, the scatter tends to increase with increasing formation radius. This increase is less significant for the youngest stars and the oldest stars. Interestingly the intermediate age stars show the largest scatter in their radial redistribution.

Bibliography

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 597, 21

Acharova I., Gibson B. K., Mishurov Y. N., Kovtyukh V., 2013, Astronomy & Astrophysics, 557, A107

Agertz O., et al., 2021, MNRAS, 503, 5826

Ahumada R., et al., 2020, ApJS, 249, 3

Anders F., et al., 2014, A&A, 564, A115

Anders F., et al., 2017, A&A, 600, A70

Andrews B. H., Martini P., 2013, ApJ, 765, 140

Andrievsky S. M., Luck R. E., Martin P., Lépine J. R. D., 2004, A&A, 413, 159

Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017, MNRAS, 470, 4698

Arellano-Córdova K. Z., Esteban C., García-Rojas J., Méndez-Delgado J. E., 2020, Monthly Notices of the Royal Astronomical Society

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, AR&AA, 47, 481

Astropy Collaboration et al., 2013, A&A, 558, A33

Baba J., Morokuma-Matsui K., Miyamoto Y., Egusa F., Kuno N., 2016, MNRAS, 460, 2472

Balser D. S., Rood R. T., Bania T. M., Anderson L. D., 2011, ApJ, 738, 27

Balser D. S., Wenger T. V., Anderson L. D., Bania T. M., 2015, ApJ, 806, 199

Belfiore F., et al., 2017, MNRAS, 469, 151

Bellardini M. A., Wetzel A., Loebman S. R., Faucher-Giguère C.-A., Ma X., Feldmann R., 2021, MNRAS, 505, 4586

Bellardini M. A., Wetzel A., Loebman S. R., Bailin J., 2022, MNRAS, 514, 4270

Benincasa S. M., et al., 2020, MNRAS, 497, 3993

Bensby T., et al., 2017, A&A, 605, A89

Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, ApJ, 773, 43

Bird J. C., Loebman S. R., Weinberg D. H., Brooks A. M., Quinn T. R., Christensen C. R., 2021, MNRAS, 503, 1815

Bland-Hawthorn J., Gerhard O., 2016, AR&AA, 54, 529

Bland-Hawthorn J., Krumholz M. R., Freeman K., 2010, ApJ, 713, 166

Bland-Hawthorn J., et al., 2019, MNRAS, 486, 1167

Boardman N., et al., 2020, MNRAS, 491, 3672

Boeche C., et al., 2013, A&A, 559, A59

Boeche C., et al., 2014, A&A, 568, A71

Bovy J., 2016, ApJ, 817, 49

Brook C. B., Kawata D., Gibson B. K., Freeman K. C., 2004, ApJ, 612, 894

Brunetti M., Chiappini C., Pfenniger D., 2011, A&A, 534, A75

Buder S., et al., 2018, MNRAS, 478, 4513

Buder S., et al., 2021, MNRAS, 506, 150

Buder S., et al., 2022, MNRAS, 510, 2407

Cantat-Gaudin T., Donati P., Vallenari A., Sordo R., Bragaglia A., Magrini L., 2016, A&A, 588, A120

Cantat-Gaudin T., et al., 2020, A&A, 640, A1

Carr C., Johnston K. V., Laporte C. F. P., Ness M. K., 2022, MNRAS, 516, 5067

Carrell K., Chen Y., Zhao G., 2012, AJ, 144, 185

Carrera R., Pancino E., 2011, A&A, 535, A30

Carrillo A., Ness M. K., Hawkins K., Sanderson R. E., Wang K., Wetzel A., Bellardini M. A., 2023, ApJ, 942, 35

Carton D., et al., 2018, Monthly Notices of the Royal Astronomical Society, 478, 4293

Casagrande L., Schönrich R., Asplund M., Cassisi S., Ramírez I., Meléndez J., Bensby T., Feltzing S., 2011, A&A, 530, A138+

Casamiquela L., Castro-Ginard A., Anders F., Soubiran C., 2021, A&A, 654, A151

Cedrés B., Cepa J., 2002, A&A, 391, 809

Cheng J. Y., et al., 2012, ApJ, 746, 149

Colbrook M. J., Ma X., Hopkins P. F., Squire J., 2017, MNRAS, 467, 2421

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781

Cui X.-Q., et al., 2012, Research in Astronomy and Astrophysics, 12, 1197

Cunha K., et al., 2016, Astronomische Nachrichten, 337, 922

Curti M., et al., 2020, Monthly Notices of the Royal Astronomical Society, 492, 821

Dalton G., et al., 2012, in Ground-based and Airborne Instrumentation for Astronomy IV. p. 84460P, doi:10.1117/12.925950

Daniel K. J., Wyse R. F. G., 2015, MNRAS, 447, 3576

Daniel K. J., Schaffner D. A., McCluskey F., Fiedler Kawaguchi C., Loebman S., 2019, ApJ, 882, 111

Davies B., Origlia L., Kudritzki R.-P., Figer D. F., Rich R. M., Najarro F., Negueruela I., Clark J. S., 2009, ApJ, 696, 2014

De Silva G. M., Freeman K. C., Bland-Hawthorn J., Asplund M., Bessell M. S., 2007, AJ, 133, 694

De Silva G. M., et al., 2015, MNRAS, 449, 2604

Dehnen W., Binney J. J., 1998, MNRAS, 298, 387

Di Matteo P., 2016, PASA, 33, e027

Donor J., et al., 2018, AJ, 156, 142

Donor J., et al., 2020, AJ, 159, 199

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., Price-Whelan A. M., Mészáros S., Nitschelm C., 2022, ApJ, 928, 23

El-Badry K., Wetzel A., Geha M., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2016, ApJ, 820, 131

El-Badry K., et al., 2018, MNRAS, 473, 1930

Escala I., et al., 2018, MNRAS, 474, 2194

Esteban C., García-Rojas J., 2018, MNRAS, 478, 2315

Esteban C., Carigi L., Copetti M. V. F., García-Rojas J., Mesa-Delgado A., Castañeda H. O., Péquignot D., 2013, MNRAS, 433, 382

Esteban C., Fang X., García-Rojas J., Toribio San Cipriano L., 2017, MNRAS, 471, 987

Faucher-Giguère C.-A., 2018, MNRAS, 473, 3717

Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, ApJ, 703, 1416

Feltzing S., Bowers J. B., Agertz O., 2020, MNRAS, 493, 1419

Fernández-Martín A., Pérez-Montero E., Vílchez J. M., Mampaso A., 2017, A&A, 597, A84

Feuillet D. K., et al., 2018, MNRAS, 477, 2326

Frankel N., Rix H.-W., Ting Y.-S., Ness M., Hogg D. W., 2018, ApJ, 865, 96

Frankel N., Sanders J., Rix H.-W., Ting Y.-S., Ness M., 2019, ApJ, 884, 99

Frankel N., Sanders J., Ting Y.-S., Rix H.-W., 2020, ApJ, 896, 15

Freeman K., Bland-Hawthorn J., 2002, AR&AA, 40, 487

Frinchaboy P. M., et al., 2013, ApJL, 777, L1

Gandhi P. J., Wetzel A., Hopkins P. F., Shappee B. J., Wheeler C., Faucher-Giguère C.-A., 2022, MNRAS, 516, 1941

Garrison-Kimmel S., Boylan-Kolchin M., Bullock J. S., Lee K., 2014, MNRAS, 438, 2578

Garrison-Kimmel S., et al., 2017, MNRAS, 471, 1709

Garrison-Kimmel S., et al., 2018, MNRAS, 481, 4133

Garrison-Kimmel S., et al., 2019a, MNRAS, 487, 1380

Garrison-Kimmel S., et al., 2019b, MNRAS, 489, 4574

Genovali K., et al., 2014, A&A, 566, A37

Gibson B. K., Pilkington K., Brook C. B., Stinson G. S., Bailin J., 2013, A&A, 554, 47

Gilmore G., et al., 2012, The Messenger, 147, 25

Glazebrook K., 2013, PASA, 30, 56

Grand R. J. J., Kawata D., Cropper M., 2012, MNRAS, 421, 1529

Grand R. J. J., Kawata D., Cropper M., 2015, MNRAS, 447, 4018

Grand R. J. J., et al., 2016, MNRAS, 460, L94

Gurvich A. B., et al., 2023, MNRAS, 519, 2598

Hafen Z., et al., 2022, MNRAS, 514, 5056

Hahn O., Abel T., 2011, MNRAS, 415, 2101

Halle A., Di Matteo P., Haywood M., Combes F., 2015, A&A, 578, A58

Halle A., Di Matteo P., Haywood M., Combes F., 2018, A&A, 616, A86

Harris C. R., et al., 2020, Nature, 585, 357

Hayden M. R., et al., 2014, AJ, 147, 116

Hayden M. R., et al., 2015, ApJ, 808, 132

Haywood M., 2008, MNRAS, 388, 1175

Hemler Z. S., et al., 2021, MNRAS, 506, 3024

Hernandez S., et al., 2021, ApJ, 908, 226

Ho I.-T., et al., 2017, The Astrophysical Journal, 846, 39

Ho I.-T., et al., 2018, A&A, 615

van den Hoek L. B., Groenewegen M. A. T., 1997, A&AS, 123, 305

Hogg D. W., et al., 2016, ApJ, 833, 262

Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins P. F., et al., 2018, MNRAS, 480, 800

Hopkins P. F., et al., 2023, MNRAS, 519, 3154

Hu G., Shao Z., Li L., 2023, arXiv e-prints, p. arXiv:2304.12929

Hunter J. D., 2007, Computing in Science Engineering, 9, 90

Iwamoto K., Brachwitz F., Nomoto K., Kishimoto N., Umeda H., Hix W. R., Thielemann F.-K., 1999, ApJS, 125, 439

Izzard R. G., Tout C. A., Karakas A. I., Pols O. R., 2004, MNRAS, 350, 407

Jones T., Ellis R., Jullo E., Richard J., 2010, ApJL, 725, L176

Jones T., Ellis R. S., Richard J., Jullo E., 2013, The Astrophysical Journal, 765, 48

Jones T., et al., 2015, The Astronomical Journal, 149, 107

de Jong R. S., et al., 2019, The Messenger, 175, 3

Jönsson H., et al., 2020, AJ, 160, 120

Kassin S. A., et al., 2012, ApJ, 758, 106

Katz D., Soubiran C., Cayrel R., Barbuy B., Friel E., Bienaymé O., Perrin M.-N., 2011, A&A, 525, A90

Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201

Kawata D., Hunt J. A. S., Grand R. J. J., Pasetto S., Cropper M., 2014, MNRAS, 443, 2757

Keller B. W., Wadsley J. W., Wang L., Kruijssen J. M. D., 2019, MNRAS, 482, 2244

Khoperskov S., Di Matteo P., Haywood M., Gómez A., Snaith O. N., 2020, A&A, 638, A144

Khoperskov S., Sivkova E., Saburova A., Vasiliev E., Shustov B., Minchev I., Walcher C. J., 2023, A&A, 671, A56

Kobayashi C., Karakas A. I., Lugaro M., 2020, ApJ, 900, 179

Kollmeier J. A., et al., 2017, arXiv e-prints, p. arXiv:1711.03234

Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019, A&A, 631, L9

Kordopatis G., et al., 2015, MNRAS, 447, 3526

Korotin S. A., Andrievsky S. M., Luck R. E., Lépine J. R. D., Maciel W. J., Kovtyukh V. V., 2014, MNRAS, 444, 3301

Kovtyukh V., et al., 2022, MNRAS, 510, 1894

Kreckel K., et al., 2019, ApJ, 887, 80

Kreckel K., et al., 2020, MNRAS, 499, 193

Kroupa P., 2001, MNRAS, 322, 231

Krumholz M. R., Gnedin N. Y., 2011, ApJ, 729, 36

Krumholz M. R., Ting Y.-S., 2018, MNRAS, 475, 2236

Leitherer C., et al., 1999, ApJS, 123, 3

Lemasle B., François P., Piersimoni A., Pedicelli S., Bono G., Laney C. D., Primas F., Romaniello M., 2008, A&A, 490, 613

Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, A&A, 500, 145

Li Z., et al., 2023, MNRAS, 518, 286

Lian J., et al., 2022, MNRAS, 511, 5639

Limongi M., Chieffi A., 2018, ApJS, 237, 13

Loebman S. R., Roškar R., Debattista V. P., Ivezić Ž., Quinn T. R., Wadsley J., 2011, ApJ, 737, 8

Loebman S. R., Debattista V. P., Nidever D. L., Hayden M. R., Holtzman J. A., Clarke A. J., Roškar R., Valluri M., 2016, ApJL, 818, L6

Lu Y., Buck T., Minchev I., Ness M. K., 2022, MNRAS, 515, L34

Luck R. E., Lambert D. L., 2011, AJ, 142, 136

Luck R. E., Kovtyukh V. V., Andrievsky S. M., 2006, AJ, 132, 902

Luck R. E., Andrievsky S. M., Kovtyukh V. V., Gieren W., Graczyk D., 2011, AJ, 142, 51

Lynden-Bell D., Kalnajs A. J., 1972, MNRAS, 157, 1

Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140

Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017a, Monthly Notices of the Royal Astronomical Society, 466, 4780

Ma X., Hopkins P. F., Wetzel A. R., Kirby E. N., Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Quataert E., 2017b, MNRAS, 467, 2430

Maciel W. J., Andrievsky S., 2019, arXiv e-prints, p. arXiv:1906.01686

Magrini L., Sestito P., Randich S., Galli D., 2009, A&A, 494, 95

Majewski S. R., et al., 2017, AJ, 154, 94

Mannucci F., Della Valle M., Panagia N., 2006, MNRAS, 370, 773

Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115

Maoz D., Graur O., 2017, ApJ, 848, 25

Marigo P., 2001, A&A, 370, 194

Marquez A., Schuster W. J., 1994, A&AS, 108, 341

Martinez-Medina L. A., Pichardo B., Moreno E., Peimbert A., 2016, MNRAS, 463, 459

Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4

Matteucci F., Francois P., 1989, MNRAS, 239, 885

McCluskey F., Wetzel A., Loebman S. R., Moreno J., Faucher-Giguere C.-A., 2023, arXiv e-prints, p. arXiv:2303.14210

Mercado F. J., et al., 2021, MNRAS, 501, 5121

Metha B., Trenti M., Chu T., Battisti A., 2022, MNRAS, 514, 4465

Miglio A., et al., 2021, A&A, 645, A85

Mikolaitis Š., et al., 2014, A&A, 572, A33

Minchev I., Famaey B., 2010, ApJ, 722, 112

Minchev I., Famaey B., Combes F., Di Matteo P., Mouhcine M., Wozniak H., 2011, A&A, 527, A147+

Minchev I., Famaey B., Quillen A. C., Di Matteo P., Combes F., Vlajić M., Erwin P., Bland-Hawthorn J., 2012, A&A, 548, A126

Minchev I., et al., 2018, MNRAS, 481, 1645

Mollá M., Díaz Á. I., Cavichia O., Gibson B. K., Maciel W. J., Costa R. D. D., Ascasibar Y., Few C. G., 2019a, MNRAS, 482, 3071

Mollá M., et al., 2019b, MNRAS, 490, 665

Muley D. A., Wheeler C. R., Hopkins P. F., Wetzel A., Emerick A., Kereš D., 2021, MNRAS, 508, 508

Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, MNRAS, 454, 2691

Muratov A. L., et al., 2017, MNRAS, 468, 4170

Naidu R. P., Conroy C., Bonaca A., Johnson B. D., Ting Y.-S., Caldwell N., Zaritsky D., Cargile P. A., 2020, ApJ, 901, 48

Ness M. K., Wheeler A. J., McKinnon K., Horta D., Casey A. R., Cunningham E. C., Price-Whelan A. M., 2022, ApJ, 926, 144

Netopil M., Paunzen E., Heiter U., Soubiran C., 2016, A&A, 585, A150

Netopil M., Oralhan İ. A., Çakmak H., Michel R., Karataş Y., 2022, MNRAS, 509, 421

Nieva M. F., Przybilla N., 2012, A&A, 539, A143

Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, Nucl. Phys. A, 777, 424

Nomoto K., Kobayashi C., Tominaga N., 2013, AR&AA, 51, 457

Okalidis P., Grand R. J. J., Yates R. M., Springel V., 2022, MNRAS, 514, 5085

Orr M. E., et al., 2023, MNRAS, 521, 3708

Ostdiek B., et al., 2020, A&A, 636, A75

Ostriker J. P., Tremaine S. D., 1975, ApJL, 202, L113

Pancino E., Carrera R., Rossetti E., Gallart C., 2010, A&A, 511, A56

Patrício V., et al., 2019, MNRAS, 489, 224

Pedicelli S., et al., 2009, A&A, 504, 81

Pignatari M., et al., 2016, ApJS, 225, 24

Pillepich A., et al., 2019, MNRAS, 490, 3196

Pilyugin L. S., Grebel E. K., Kniazev A. Y., 2014, AJ, 147, 131

Planck Collaboration et al., 2020, A&A, 641, A6

Poetrodjojo H., et al., 2018, MNRAS, 479, 5235

Porter L. E., Orr M. E., Burkhart B., Wetzel A., Ma X., Hopkins P. F., Emerick A., 2022, MNRAS, 515, 3555

Prantzos N., Boissier S., 2000, Monthly Notices of the Royal Astronomical Society, 313, 338

Press W. H., Schechter P., 1974, ApJ, 187, 425

Price-Jones N., et al., 2020, Monthly Notices of the Royal Astronomical Society, 496, 5101

Price-Whelan A. M., et al., 2018, AJ, 156, 123

Queyrel J., et al., 2012, A&A, 539, 93

Quillen A. C., Minchev I., Bland-Hawthorn J., Haywood M., 2009, MNRAS, 397, 1599

Quillen A. C., Nolting E., Minchev I., De Silva G., Chiappini C., 2018, MNRAS, 475, 4450

Ratcliffe B. L., Ness M. K., Buck T., Johnston K. V., Sen B., Beraldo e Silva L., Debattista V. P., 2022, ApJ, 924, 60

Reddy A. B. S., Giridhar S., Lambert D. L., 2020, Journal of Astrophysics and Astronomy, 41, 38

Rennehan D., Babul A., Hopkins P. F., Davé R., Moa B., 2019, MNRAS, 483, 3810

Rohatgi A., 2022, Webplotdigitizer: Version 4.6, `https://automeris.io/WebPlotDigitizer`

Roškar R., Debattista V. P., Stinson G. S., Quinn T. R., Kaufmann T., Wadsley J., 2008a, ApJL, 675, L65

Roškar R., Debattista V. P., Quinn T. R., Stinson G. S., Wadsley J., 2008b, ApJL, 684, L79

Rudolph A. L., Fich M., Bell G. R., Norsen T., Simpson J. P., Haas M. R., Erickson E. F., 2006, ApJS, 162, 346

Ryden B., 2017, Introduction to cosmology. Cambridge University Press

Sakhibov F., Zinchenko I. A., Pilyugin L. S., Grebel E. K., Just A., Vílchez J. M., 2018, MNRAS, 474, 1657

Sánchez-Menguiano L., et al., 2016, A&A, 587, A70

Sánchez-Menguiano L., Sánchez S. F., Pérez I., Ruiz-Lara T., Galbany L., Anderson J. P., Kuncarayakti H., 2020, MNRAS, 492, 4149

Sánchez S. F., et al., 2015, A&A, 573, 105

Sanders J. L., Binney J., 2016, MNRAS, 457, 2107

Sanders N. E., Caldwell N., McDowell J., Harding P., 2012, ApJ, 758, 133

Sanderson R. E., et al., 2020, ApJS, 246, 6

Santistevan I. B., Wetzel A., El-Badry K., Bland-Hawthorn J., Boylan-Kolchin M., Bailin J., Faucher-Giguère C.-A., Benincasa S., 2020, MNRAS, 497, 747

Santos-Peral P., Recio-Blanco A., Kordopatis G., Fernández-Alvar E., de Laverny P., 2021, A&A, 653, A85

Schönrich R., Binney J., 2009, MNRAS, 396, 203

Searle L., 1977, in Tinsley B. M., Larson Richard B. Gehret D. C., eds, Evolution of Galaxies and Stellar Populations. p. 219

Searle L., Zinn R., 1978, ApJ, 225, 357

Sellwood J. A., Binney J. J., 2002, MNRAS, 336, 785

Sestito P., Bragaglia A., Randich S., Pallavicini R., Andrievsky S. M., Korotin S. A., 2008, A&A, 488, 943

Sharda P., Krumholz M. R., Wisnioski E., Forbes J. C., Federrath C., Acharyya A., 2021, MNRAS, 502, 5935

Sharda P., Ginzburg O., Krumholz M. R., Forbes J. C., Wisnioski E., Mingozzi M., Zovaro H. R. M., Dekel A., 2023, arXiv e-prints, p. arXiv:2303.15853

Sharma S., Hayden M. R., Bland-Hawthorn J., 2021, MNRAS, 507, 5882

Solar M., Tissera P. B., Hernandez-Jimenez J. A., 2020, MNRAS, 491, 4894

Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2017, MNRAS, 466, 88

Spina L., et al., 2017, A&A, 601, A70

Spina L., et al., 2021, MNRAS, 503, 3279

Spitoni E., Cescutti G., Minchev I., Matteucci F., Silva Aguirre V., Martig M., Bono G., Chiappini C., 2019, A&A, 628, A38

Stanghellini L., Haywood M., 2010, ApJ, 714, 1096

Stanghellini L., Haywood M., 2018, ApJ, 862, 45

Stanghellini L., Magrini L., Villaver E., Galli D., 2010, A&A, 521, A3

Stanghellini L., Magrini L., Casasola V., Villaver E., 2014, A&A, 567, A88

Steinmetz M., et al., 2006, AJ, 132, 1645

Stern J., et al., 2021, ApJ, 911, 88

Stott J. P., et al., 2014, MNRAS, 443, 2695

Su K.-Y., Hopkins P. F., Hayward C. C., Faucher-Giguère C.-A., Kereš D., Ma X., Robles V. H., 2017, MNRAS, 471, 144

Sukhbold T., Ertl T., Woosley S. E., Brown J. M., Janka H. T., 2016, ApJ, 821, 38

Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., McCarthy I. G., Crain R. A., Theuns T., 2012, MNRAS, 426, 935

Swinbank A. M., et al., 2017, MNRAS, 467, 3140

Takada M., et al., 2014, PASJ, 66, R1

Tanabashi M., et al., 2018, Phys. Rev. D, 98, 030001

The MSE Science Team et al., 2019, arXiv e-prints, p. arXiv:1904.04907

Ting Y.-S., Weinberg D. H., 2022, ApJ, 927, 209

Ting Y.-S., De Silva G. M., Freeman K. C., Parker S. J., 2012, MNRAS, 427, 882

Tissera P. B., Rosas-Guevara Y., Bower R. G., Crain R. A., Lagos C. d. P., Schaller M., Schaye J., Theuns T., 2019, Monthly Notices of the Royal Astronomical Society, 482, 2208

Tremonti C. A., et al., 2004, ApJ, 613, 898

Tronrud T., Tissera P. B., Gómez F. A., Grand R. J. J., Pakmor R., Marinacci F., Simpson C. M., 2022, MNRAS, 515, 3818

Übler H., et al., 2019, ApJ, 880, 48

Vera-Ciro C., D'Onghia E., Navarro J., Abadi M., 2014, ApJ, 794, 173

Vickers J. J., Shen J., Li Z.-Y., 2021, ApJ, 922, 189

Vincenzo F., Kobayashi C., 2020, MNRAS, 496, 80

Virtanen P., et al., 2020, Nature Methods, 17, 261

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Vogt F. P. A., Pérez E., Dopita M. A., Verdes-Montenegro L., Borthakur S., 2017, A&A, 601, 61

Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83

Wang C., et al., 2019, MNRAS, 482, 2189

Wang X., et al., 2020, ApJ, 900, 183

Wellons S., et al., 2023, MNRAS, 520, 5394

Wenger T. V., Balser D. S., Anderson L. D., Bania T. M., 2019, ApJ, 887, 114

Wetzel A., Garrison-Kimmel S., 2020, GizmoAnalysis: Read and analyze Gizmo simulations (ascl:2002.015)

Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJL, 827, L23

Wetzel A., et al., 2023, ApJS, 265, 44

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574

Wisnioski E., et al., 2015, ApJ, 799, 209

Wojno J., et al., 2016, MNRAS, 461, 4246

Wuyts E., et al., 2016, ApJ, 827, 74

Xiang M.-S., et al., 2015, Research in Astronomy and Astrophysics, 15, 1209

Xiang M., et al., 2017, ApJS, 232, 2

Yu S., et al., 2020, MNRAS, 494, 1539

Yu S., et al., 2021, MNRAS, 505, 889

Yu S., et al., 2022, arXiv e-prints, p. arXiv:2210.03845

Yuan T. T., Kewley L. J., Rich J., 2013, ApJ, 767, 106

Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, ApJL, 771, L19

Zhang H., Chen Y., Zhao G., 2021, ApJ, 919, 52

Zinchenko I. A., Pilyugin L. S., Grebel E. K., Sánchez S. F., Vílchez J. M., 2016, MNRAS, 462, 2715

Zinchenko I., Just A., Pilyugin L., Lara-Lopez M., 2019, Astronomy & Astrophysics, 623, A7

Zurita A., Bresolin F., 2012, MNRAS, 427, 1463