Decoding Galaxy Evolution with Gas-phase and Stellar Elemental Abundances

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Brett Hon Wing Kao Andrews

Graduate Program in Astronomy

The Ohio State University

2014

Dissertation Committee:
Professor David H. Weinberg, Advisor
Professor Jennifer A. Johnson
Professor Marc H. Pinsonneault
Copyright by

Brett Hon Wing Kao Andrews

2014
Abstract

Elemental abundances of gas and stars are sensitive diagnostics of the main processes that drive galaxy evolution: gas inflow, star formation, enrichment, and gas outflow. The relation between galaxy stellar mass and gas-phase oxygen abundance, known as the mass–metallicity relation (MZR), is one of the strongest constraints on galaxy evolution models. However, the popular strong line methods of measuring oxygen abundance have large systematic uncertainties. We employ a more robust direct method to measure the metallicities of ∼200,000 star-forming galaxies from the Sloan Digital Sky Survey stacked in bins of stellar mass and star formation rate (SFR) to significantly enhance the signal-to-noise ratio of the weak auroral lines required for the direct method. The direct method MZR has a steeper slope, a lower turnover mass, and a factor of 2–3 greater dependence on the SFR than strong line MZRs.

Gas-phase abundances reflect a galaxy’s current abundances, but stellar abundances encode its complete enrichment history. We use a sample of 35 microlensed bulge dwarf and subgiant stars, whose brightness increased by a factor of 100–1000 due to lensing from an intervening star, to investigate the formation of the Galactic bulge. We apply principal component abundance analysis (PCAA)—a
principal component decomposition of relative abundances [X/Fe]—to this sample to characterize its distribution in the 12-dimensional space defined by their elemental abundances. The first principal component PC1, which describes the abundance patterns of most stars in the sample, shows a strong contribution from $\alpha$-elements, reflecting the relative contributions of core-collapse and Type Ia supernovae. The second principal component PC2 is characterized by a Na–Ni correlation, the likely product of metallicity-dependent core-collapse supernova (CCSN) yields. Applying PCAA to a sample of local disk dwarfs yields a nearly identical PC1, suggesting broadly similar $\alpha$-enrichment histories. However, the disk PC2 is dominated by a Y–Ba correlation, likely indicating a contribution of $s$-process enrichment from long-lived asymptotic giant branch stars that is absent from the bulge PC2 because of rapid bulge formation.

Detailed interpretations of stellar abundance patterns require chemical evolution modeling, but these models are sensitive to the treatment of star formation efficiency (SFE), outflow, stellar yields, and mixing of stellar populations. The two main features in [$\alpha$/Fe]–[Fe/H] are the [Fe/H] of the knee (where the high [$\alpha$/Fe] plateau turns downwards towards solar ratios), and the equilibrium abundance to which the simulations quickly asymptote. A higher SFE increases the [Fe/H] of the knee but does not change the equilibrium abundance. Conversely, a higher outflow mass-loading parameter does not impact the [Fe/H] of the knee, but it decreases the equilibrium [Fe/H]. Adopting different CCSN yields makes a modest impact on
the evolution of $\alpha$-elements and Fe but produces a dramatically different evolution for elements with a strong metallicity dependence to their yields, like Na and Al.

The most natural explanation for reproducing the bimodality in $[\alpha/\text{Fe}]-[\text{Fe}/\text{H}]$ involves mixing stellar populations born at different galactocentric radii with unique enrichment histories.
Dedication

For my parents, my sister, and Lorrie.
Acknowledgments

I can only begin to express my deep appreciation for David Weinberg’s superb guidance that made this dissertation possible. His excellent grasp of the larger picture continually forced me to understand our results in a broader context. His questions cut straight to the heart of the matter and would evoke sheer terror if asked by anyone who does not possess David’s constantly optimistic outlook. I have benefitted enormously from such insightful (and cheerful) interrogations, and he has instilled in me an appreciation for understanding the failure modes of models, which are often more interesting than the successes. It is this thought process that I try to emulate and nowhere is it more apparent than in David’s eloquent writing and precise editing. I quickly realized that attempting to improve text that he has written is an exercise in frustration. I could not have asked for more in a mentor and feel lucky to have had the honor and joy of working with him.

Most graduate students are fortunate to have one great advisor, but I hit the jackpot and ended up with two. Both David and I would have been lost without Jennifer Johnson’s expertise in chemical evolution. She selflessly shared her deep knowledge of stellar abundances and spectroscopy, and she has helped me learn how to recognize spurious observational trends. She is always available for advice and

vi
will drop everything to help. She is especially talented at inspiring her students to take initiative, and her belief in me has always bolstered my self-confidence. I will sorely miss interacting with Jennifer on a daily basis.

In addition to my advisors, I have worked closely with a few other faculty members who have significantly aided in my development as a scientist. I would like to thank Marc Pinsonneault for serving on my dissertation committee and sharing his preternatural understanding of stellar interiors that frequently clarified my confusion about the inner workings of stars. I am indebted to Paul Martini for his patient mentorship as he taught me the subtleties of spectroscopy, which resulted in Chapter 2 of this dissertation. Todd Thompson has my endless appreciation for advising my first project in graduate school and instilling an atmosphere of excitement about astronomy that is dangerously contagious.

A large part of what makes Ohio State a special place is its interactive environment, including the daily Coffee discussions and open door philosophy. I have enjoyed many conversations with other faculty members at Ohio State, especially Rick Pogge, Chris Kochanek, Andy Gould, and Scott Gaudi for their assistance on my research projects and job applications.

I truly appreciate all that I have learned from other Ohio State graduate students and postdocs, whose genuine interest in my research struggles made it possible to overcome many obstacles. In particular, I recognize the assistance of
Jon Bird, Molly Peeples, Ralph Schönrich, Katie Schlesinger, Courtney Epstein, Jill Gerke, and Ben Shappee, though I have benefited from the support of nearly all of the graduate students with which I have overlapped. I would like to express a special thanks to a wonderful, smart, and witty group of officemates, including Linda Watson, Tom Beatty, Jen Van Saders, and Matt Penny—your positive effect on my daily happiness has made this journey much more entertaining.

I am grateful to Danilo Marchesini and Pieter van Dokkum for a welcoming introduction into astronomy research as an undergraduate and to Meg Urry and Charles Bailyn for sage advice that helped launch my astronomy career.

I could not have achieved nearly as much without the camaraderie of my roommates, Travis Nelson and Calen Henderson, and my best friend, William Stewart.

My whole life I have been fortunate to have outstanding parents—though I only recognized it recently. They provided consistent encouragement and opportunities from day one that gave me the opportunity to thrive. My parents say that younger sister admired me growing up, but it is I who look up to her now. I am incredibly proud of her achievements and appreciate her neverending support, especially since she started her graduate career and has given me suggestions about succeeding in graduate school.
Most of all, I am forever grateful to my wife, whose continual love, frequent sacrifices, and thoughtful advice have carried me throughout this journey.
Vita

June 6, 1986 ................. Born – Hamden, Connecticut, USA

2008 ......................... B.S. Astronomy & Physics, Yale University

2008 – 2014 .................... Graduate Teaching and Research Associate,

The Ohio State University

Publications

Research Publications


Fields of Study

Major Field: Astronomy
# Table of Contents

Abstract ........................................................................ ii

Dedication ..................................................................... v

Acknowledgments .......................................................... vi

Vita .............................................................................. x

List of Figures ................................................................. xvii

Chapter 1: Introduction .................................................. 1

1.1 Chemical Evolution of Galaxies ............................... 1

1.2 Relation to Previous Work .................................... 3

1.3 Scope of the Dissertation ...................................... 6

Chapter 2: The Mass–Metallicity Relation with the Direct Method on Stacked Spectra of SDSS Galaxies .................... 9

2.1 Introduction .............................................................. 9

2.2 Method ..................................................................... 17

2.2.1 Sample Selection .................................................. 17

2.2.2 Stacking Procedure ............................................. 19

2.2.3 Stellar Continuum Subtraction ............................ 21

2.2.4 Automated Line Flux Measurements .................... 23

2.3 Electron Temperature and Direct Abundance Determination ........................................... 25

2.3.1 Electron Temperatures ....................................... 25

2.3.2 Ionic and Total Abundances ............................... 30
### 2.3.3 Strong Line Metallicities

- 34

### 2.4 How Does Stacking Affect Measured Electron Temperatures and Metallicities?

- 35

### 2.5 The Mass–Metallicity Relation and Mass–Metallicity–SFR Relation

- 40
  - 2.5.1 The Mass–Metallicity Relation
  - 40
  - 2.5.2 Mass–Metallicity–SFR Relation
  - 43
  - 2.5.3 The Fundamental Metallicity Relation
  - 47

### 2.6 N/O Abundance

- 50

### 2.7 Discussion

- 55
  - 2.7.1 Comparison to a Previous Analysis That Used Auroral Lines from Stacked Spectra
  - 55
  - 2.7.2 Temperature and Metallicity Discrepancies
  - 58
  - 2.7.3 Strong Line Calibrations and the SFR-dependence of the FMR
  - 60
  - 2.7.4 Physical Processes Governing the MZR and $M_\star – Z – SFR$ Relation
  - 66

### 2.8 Summary

- 70

### 2.9 Figures

- 73

---

### Chapter 3: Principal Component Abundance Analysis of Microlensed Bulge Dwarf and Subgiant Stars

- 89

#### 3.1 Introduction

- 89

#### 3.2 Method

- 91

#### 3.3 Results

- 92

#### 3.4 Conclusions

- 99

#### 3.5 Figures

- 102

---

### Chapter 4: Inflow, Outflow, Yields, and Stellar Mixing in Chemical Evolution Models

- 106

#### 4.1 Introduction

- 106

---

xv
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Distribution of our sample of SDSS galaxies in $M_*$ and SFR</td>
<td>74</td>
</tr>
<tr>
<td>2.2</td>
<td>Example stacked galaxy spectra</td>
<td>74</td>
</tr>
<tr>
<td>2.3</td>
<td>$T_e[\text{O<del>II}], T_e[\text{N</del>II}], \text{and } T_e[\text{S<del>II}] \text{ vs. } T_e[\text{O</del>III}]$ for the galaxy stacks</td>
<td>75</td>
</tr>
<tr>
<td>2.4</td>
<td>$T_e[\text{O<del>II}] - T_e[\text{N</del>II}]$ relation for the galaxy stacks</td>
<td>76</td>
</tr>
<tr>
<td>2.5</td>
<td>$O^+/H$ and $O^{++}/H$ ionic abundances of the galaxy stacks</td>
<td>77</td>
</tr>
<tr>
<td>2.6</td>
<td>Self-calibration of the $T_2 - T_3$ relation and its effect on the mass–metallicity relation</td>
<td>78</td>
</tr>
<tr>
<td>2.7</td>
<td>Comparison of $T_e[\text{O<del>III}], T_e[\text{O</del>II}], \text{and direct method metallicities}$ between individual galaxies and galaxy stacks</td>
<td>79</td>
</tr>
<tr>
<td>2.8</td>
<td>$T_e[\text{O<del>II}], T_e[\text{N</del>II}], \text{and } T_e[\text{S<del>II}] \text{ vs. } T_e[\text{O</del>III}]$ for individual galaxies and galaxy stacks</td>
<td>80</td>
</tr>
<tr>
<td>2.9</td>
<td>$[\text{O<del>II}] \lambda 3727/H\beta$ and $[\text{O</del>III}] \lambda 5007/H\beta$ of individual galaxies and a galaxy stack</td>
<td>81</td>
</tr>
<tr>
<td>2.10</td>
<td>Direct method mass–metallicity relation</td>
<td>82</td>
</tr>
<tr>
<td>2.11</td>
<td>Direct method mass–metallicity–SFR relation</td>
<td>83</td>
</tr>
<tr>
<td>2.12</td>
<td>Direct method fundamental metallicity relation</td>
<td>84</td>
</tr>
<tr>
<td>2.13</td>
<td>Mass–metallicity–SFR for two strong line metallicity calibrations</td>
<td>85</td>
</tr>
<tr>
<td>2.14</td>
<td>$N/O$ vs. $O/H$ and $M_*$</td>
<td>86</td>
</tr>
<tr>
<td>2.15</td>
<td>Comparison between our direct method mass–metallicity relation and that of Liang et al. (2007)</td>
<td>87</td>
</tr>
<tr>
<td>2.16</td>
<td>Excitation parameter vs. $R_{23}$ and $[\text{O<del>II}] \lambda 3727/H\beta$ vs. $[\text{O</del>III}] \lambda 5007/H\beta$</td>
<td>88</td>
</tr>
</tbody>
</table>

### Supplemental Figure

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>PC1 and PC2 of the microlensed bulge dwarfs and disk dwarfs</td>
<td>103</td>
</tr>
</tbody>
</table>
3.2 PC1 histogram and PC1–PC2 space for the microlensed bulge dwarfs
3.3 $\chi^2$-fitting of principal components to abundance patterns and $\chi^2_{\text{red}}$ histogram
4.1 [O/Fe]–[Fe/H] for a closed box, inflow-only, and inflow + outflow simulations
4.2 The mass of oxygen and iron as a function of [Fe/H]
4.3 [O/Fe]–[Fe/H] for inflow timescale and inflow rate variations
4.4 [O/Fe]–[Fe/H] for star formation efficiency and outflow mass-loading parameter variations
4.5 [O/Fe]–[Fe/H] for inflow metallicity and stellar initial mass function variations
4.6 Star formation histories for various functional forms of the inflow rate
4.7 [O/Fe]–[Fe/H] for the minimum SNIa delay time
4.8 SNIa delay time distribution and number of SNIa as a function of time
4.9 [O/Fe]–[Fe/H] for various SNIa delay time distributions
4.10 [O/Fe]–[Fe/H] for a simulation with a warm phase of the ISM
4.12 Mass of various elements and iron as a function of [Fe/H]
4.13 [O/Fe]–[Fe/H] for the radial mixing and two infall simulations
Chapter 1: Introduction

1.1. Chemical Evolution of Galaxies

Despite nearly a century of effort, our understanding of galaxy formation and evolution remains incomplete. One of the first important steps towards our current picture of galaxy formation was the early work by Oort (1926) and Baade (1944) that discovered the existence of multiple populations of stars in the Milky Way and Andromeda, respectively. O’Connell (1958) found that distinct populations formed distinct structural components of the Milky Way. The canonical paper by Eggen et al. (1962) expanded on this idea by interpreting correlations between stellar metallicities and dynamics—for example, lower metallicity stars have larger $W$ velocities, larger inferred scale heights, and larger ellipticities—as evidence for the rapid (less than a dynamical time scale) formation of the stellar halo. Their paper established the subject of “near-field cosmology” by demonstrating that stellar abundances and dynamics can be leveraged to study galaxy formation. The Eggen et al. (1962) model was later challenged by Searle & Zinn (1978), who argued that the lack of correlation between the metallicities of halo globular clusters and their galactocentric distances implied that the halo did not assemble quickly and
uniformly. Instead, they argued that smaller objects formed and merged together over a long period of time. The Searle & Zinn (1978) model is more consistent with our present model of galaxy formation, where dark matter halos form in a hierarchical process (Tinsley & Larson 1977) and gas accretes into dark matter halos. However, pinning down the details of the baryonic processes that govern galaxy formation, especially the assembly of the luminous components, has proven to be challenging. Much of the difficulty lies in accurately tracking the properties of gas due to its collisional nature and the wide range of time scales, spatial scales, temperatures, and densities that it can cover.

Fortunately, chemical abundances are sensitive to the major processes that drive galaxy evolution, namely, inflow, star formation, enrichment, and outflow. Inflow from the intergalactic medium (IGM) delivers pristine gas to galaxies, providing the raw fuel for star formation while temporarily diluting its metallicity. The ensuing star formation builds stellar mass and enriches the interstellar medium (ISM) with metals. Feedback from massive stars can drive galactic-scale outflows that eject gas and metals back into the IGM. These processes, especially inflow and outflow, are difficult to study directly but can be understood through their effects on chemical abundances. Gas-phase and stellar abundances provide complementary information on galaxy evolution because gas-phase abundances reflect recent enrichment while stellar abundances record the full enrichment history of galaxies.
1.2. Relation to Previous Work

Measuring the gas-phase abundances of large galaxy samples allows us to connect snapshots of individual objects to reconstruct an evolutionary sequence and understand how galaxy formation processes affect chemical abundances. The first indication of a correlation between mass and metallicity came when Lequeux et al. (1979) demonstrated the existence of a relation between total mass and metallicity for irregular and blue compact galaxies. Subsequent studies showed that metallicity also correlates with other galaxy properties, such as luminosity (Rubin et al. 1984) and rotation velocity (Zaritsky et al. 1994; Garnett 2002). The advent of reliable stellar population synthesis models (Bruzual & Charlot 2003) enabled more accurate stellar mass ($M_\star$) measurements from spectral energy distributions. Tremonti et al. (2004, hereafter T04) showed the existence of a tight correlation between galaxy stellar mass and metallicity among $\sim$53,000 galaxies from the Sloan Digital Sky Survey (SDSS; York et al. 2000) DR2 (Abazajian et al. 2003) based on the stellar mass measurements from Kauffmann et al. (2003a). T04 found that the scatter in the mass–metallicity (MZR) was smaller than the scatter in the luminosity–metallicity relation and concluded that the MZR was more physically motivated. Later work by Ellison et al. (2008) discovered that galaxies with high star formation rates (SFRs) are systematically offset to lower metallicities than more weakly star-forming galaxies at the same stellar mass. Mannucci et al. (2010) and Lara-López et al. (2010) studied this effect in a systematic fashion and demonstrated that the scatter
in the MZR is reduced further by accounting for SFR. Interestingly, Mannucci et al. (2010) and Lara-López et al. (2010) found that the $M_\star-Z-SFR$ relation does not evolve with redshift up to $z \sim 2.5$, as opposed to the MZR (Erb et al. 2006; Maiolino et al. 2008; Zahid et al. 2011; Moustakas et al. 2011), which shifts to lower metallicity with increasing redshift. However, the majority of these studies relied on strong line metallicity calibrations that suffer from large systematic uncertainties.

We constructed improved versions of the MZR and $M_\star-Z-SFR$ relation by stacking the spectra of $\sim 200,000$ galaxies from SDSS to detect the weak auroral lines required for the more robust direct method of measuring metallicity.

The Milky Way provides a unique case study for the chemical evolution of galaxies because its complete enrichment history is encoded in its stars (e.g., Freeman & Bland-Hawthorn 2002). Since we can resolve individual stars, we can disentangle the formation histories of its major components: the bulge, disk, and halo. Until recently, stellar abundance surveys either measured a few elemental abundances for up to $\sim 100,000$ stars (e.g., SEGUE/SEGUE-2; Yanny et al. 2009) or many ($>10$) elemental abundances for smaller samples of $100$–$1000$ stars (Bensby et al. 2003, 2005, 2014; Reddy et al. 2003, 2006; Adibekyan et al. 2012). Ongoing stellar abundance surveys such as APOGEE, Gaia-ESO, and GALAH will usher in a new generation with measurements for both many ($>10$) elements and large (100,000–1,000,000) sample sizes. These large, multi-dimensional data sets are a natural target for advanced statistical techniques, such as Principal Component
Abundance Analysis (PCAA). As a pilot study, we applied PCAA to a sample of microlensed bulge dwarf stars and a comparison sample of local disk stars each with 12 measured abundances. PCAA revealed that the primary trend of correlated α-elements is the same in both samples, but the secondary Na–Ni trend in the bulge stars compared to the Y–Ba trend in the disk stars suggests that the bulge formed more rapidly than the local disk.

Interpreting the wealth of information from these surveys requires galaxy evolution models that track chemical enrichment. No one technique currently has the power to do everything, so a range of models is needed. Smoothed particle hydrodynamics (SPH; e.g., Finlator & Davé 2008) and adaptive mesh refinement (AMR; e.g., Vogelsberger et al. 2013) models have tracked the enrichment of large populations of galaxies set in a ΛCDM Universe and have reproduced many features of the MZR. These models lack the spatial resolution and number of elements to predict stellar abundance trends. At the other end of the spectrum are classical chemical evolution models (e.g., Tinsley & Larson 1978; Matteucci & Francois 1989; Chiappini et al. 1997) that often use coupled differential equations to track the evolution of many elements in one galaxy. These models have increased in complexity by dividing the Galaxy into multiple radial zones (e.g., Tinsley 1980) and incorporating stellar dynamics such as radial migration (Schönrich & Binney 2009a).

One major advantage of classical chemical evolution models is that they are fast and flexible, so large swaths of parameter space can be explored. Recently, cosmological
zoom-in simulations (e.g., Guedes et al. 2011) have started to bridge the gap between the coarsely resolved SPH and AMR models and classical chemical evolution models because they form galaxies in a cosmological context but re-simulate them at high resolution to track detailed structure and enrichment. At present, these simulations are computationally expensive, so they are limited to only one or a few galaxies.

We have developed a one-zone classical chemical evolution model to isolate the effects of individual galaxy evolution processes and stellar yields. We investigate possible scenarios for producing scatter in abundances and apply PCAA to the model abundances to make predictions for the upcoming APOGEE data set.

1.3. Scope of the Dissertation

The MZR is one of the strongest observational constraints on galaxy evolution models. Large galaxy surveys determine metallicities with calibrations that associate the flux ratios of strong emission lines to gas-phase oxygen abundances because the auroral lines required for the more robust “direct method” are undetected in the vast majority of galaxies (>99%; Izotov et al. 2006). However, strong line calibrations suffer from large (up to 0.7 dex) systematic offsets (Kewley & Ellison 2008), which critically impacts the physical interpretation of the MZR. In Chapter 2, we stack the spectra of ~200,000 star-forming galaxies from the Sloan Digital Sky Survey (SDSS; York et al. 2000) in narrow bins of stellar mass and star formation rate (SFR) to detect the auroral lines and measure the direct method MZR and

Bulges and spheroids contain $\sim 60\%$ (Gadotti 2009) of the stellar mass in the present day Universe, so understanding bulge formation is an important aspect of any theory of galaxy formation. The proximity of the Milky Way’s bulge provides a unique opportunity to study the stellar abundances of individual stars and reconstruct its formation history. Dwarf stars are preferable to giant stars for measuring elemental abundance trends because their spectra are less challenging to analyze (Edvardsson et al. 1993). Their intrinsic faintness ($V = 19–20$; Feltzing & Gilmore 2000) makes it difficult to observe them in the bulge under normal conditions. Fortunately, gravitational microlensing by an intervening lens star can increase their brightness by a factor of 100–1000, enabling high-resolution spectroscopy that is required to determine multi-element abundances. We adopted a sample of 35 microlensed bulge dwarf and subgiant stars, most with abundances for 11 or 12 elements. In Chapter 3, we applied principal component abundance analysis (PCAA)—a principal component decomposition of stellar abundances—to this sample and a comparison sample of local disk stars to discover trends in their elemental abundances that reveal clues about the formation histories of the bulge and local disk. This chapter first appeared as Andrews et al. (2012).

While basic conclusions can be reached from by-eye analyses of elemental abundance trends, more in-depth interpretations require chemical evolution models.
The major parameters in these models are inflow rate, SFE, outflow rate, the Type Ia supernova (SNIa) delay time distribution (DTD), the stellar initial mass function (IMF), and stellar yields. Most classical chemical evolution models find one set of parameters that matches some set of observables. Unfortunately, these parameters can be partially or completely degenerate with each other. In Chapter 4, we do a systematic investigation of the effects of changing the main parameters of chemical evolution models with an emphasis on the canonical $[\text{O/Fe}]$–$[\text{Fe/H}]$ diagram. We also study how the core collapse supernova (CCSN) yields affect multi-element abundances. We conclude by applying PCAA to our simulated stellar abundances to provide guidance for a similar analysis of the upcoming APOGEE data set.
Chapter 2: The Mass–Metallicity Relation with the Direct Method on Stacked Spectra of SDSS Galaxies

2.1. Introduction

Galaxy metallicities are one of the fundamental observational quantities that provide information about their evolution. The metal content of a galaxy is governed by a complex interplay between cosmological gas inflow, metal production by stars, and gas outflow via galactic winds. Inflows dilute the metallicity of a galaxy in the short term but provide the raw fuel for star formation on longer timescales. This gas turns into stars, which convert hydrogen and helium into heavier elements. The newly formed massive stars inject energy and momentum into the gas, driving large-scale outflows that transport gas and metals out of the galaxy. The ejected metals can escape the gravitational potential well of the galaxy to enrich the intergalactic medium or reaccrete onto the galaxy and enrich the inflowing gas. This cycling of baryons in and out of galaxies directly impacts the stellar mass \( M_\star \), metallicity \( Z \), and star formation rate (SFR) of the galaxies. Thus, the galaxy stellar mass–metallicity relation (MZR) and the stellar mass–metallicity–SFR relation serve as crucial observational constraints for galaxy evolution models that
attempt to understand the build up of galaxies across cosmic time. Here we present new measurements of the MZR and the $M_* - Z - \text{SFR}$ relation that span three orders of magnitude in stellar mass with metallicities measured with the direct method.

The first indication of a correlation between mass and metallicity came when Lequeux et al. (1979) demonstrated the existence of a relation between total mass and metallicity for irregular and blue compact galaxies. Subsequent studies showed that metallicity also correlates with other galaxy properties, such as luminosity (Rubin et al. 1984) and rotation velocity (Zaritsky et al. 1994; Garnett 2002). The advent of reliable stellar population synthesis models (Bruzual & Charlot 2003) enabled more accurate stellar mass measurements from spectral energy distributions. Tremonti et al. (2004, hereafter T04) showed the existence of a tight correlation between galaxy stellar mass and metallicity among $\sim 53,000$ galaxies from the Sloan Digital Sky Survey (SDSS; York et al. 2000) DR2 (Abazajian et al. 2003) based on the stellar mass measurements from Kauffmann et al. (2003a). The T04 MZR increases as roughly $O/H \propto M_*^{1/3}$ from $M_* = 10^{8.5} - 10^{10.5} \, M_\odot$ and then flattens above $M_* \sim 10^{10.5} \, M_\odot$. They found that the scatter in the MZR was smaller than the scatter in the luminosity–metallicity relation and concluded that the MZR was more physically motivated. Lee et al. (2006) extended the MZR down another $\sim 2.5$ dex in stellar mass with a sample of local dwarf irregular galaxies. The scatter and slope of the Lee et al. (2006) MZR are consistent with the T04 MZR (cf., Zahid et al. 2012a), but the Lee et al. (2006) MZR is offset to lower metallicities by $0.2 - 0.3$
dex. This offset is likely because T04 and Lee et al. (2006) use different methods to estimate metallicity. Later work by Ellison et al. (2008) discovered that galaxies with high SFRs (and larger half-light radii) are systematically offset to lower metallicities than more weakly star-forming galaxies at the same stellar mass. Mannucci et al. (2010) and Lara-López et al. (2010) studied this effect in a systematic fashion and demonstrated that the scatter in the MZR is reduced further by accounting for SFR. Mannucci et al. (2010) introduced the concept of the fundamental metallicity relation (FMR) by parametrizing the second-order dependence of the MZR on SFR with a new abscissa,

\[
\mu_\alpha \equiv \log(M_\star) - \alpha \log(\text{SFR}),
\]

where the coefficient \(\alpha\) is chosen to minimize the scatter in the relation. We will refer to this particular parametrization as the FMR but the general relation as the \(M_\star-Z-SFR\) relation. Interestingly, Mannucci et al. (2010) and Lara-López et al. (2010) found that the \(M_\star-Z-SFR\) relation does not evolve with redshift up to \(z \sim 2.5\), as opposed to the MZR (Erb et al. 2006; Maiolino et al. 2008; Zahid et al. 2011; Moustakas et al. 2011). However, this result depends on challenging high redshift metallicity measurements, specifically the Erb et al. (2006) sample of stacked galaxy spectra at \(z \sim 2.2\) and the Maiolino et al. (2008) sample of nine galaxies at \(z \sim 3.5\).
Galaxy evolution models aim to reproduce various features of the MZR and $M_\star-Z-SFR$ relation, specifically their slope, shape, scatter, and evolution. The most distinguishing characteristic of the shape of the MZR is that it appears to flatten and become independent of mass at $M_\star \sim 10^{10.5} \, M_\odot$. The canonical explanation is that this turnover reflects the efficiency of metal ejection from galaxies because the gravitational potential wells of galaxies at and above this mass scale are too deep for supernova-driven winds to escape (Dekel & Silk 1986; Dekel & Woo 2003; Tremonti et al. 2004). In this scenario, the metallicity of these galaxies approaches the effective yield of the stellar population. However, recent simulations by Oppenheimer & Davé (2006), Finlator & Davé (2008), and Davé et al. (2011a,b) show that winds characterized by a constant velocity and constant mass-loading parameter (mass outflow rate divided by SFR; their $cw$ simulations), which were intended to represent supernova-driven winds, result in a MZR that fails to qualitatively match observations. The $cw$ simulations produce a MZR that is flat with a very large scatter at low mass, yet becomes steep above the blowout mass, which is the critical scale above which all metals are retained. Instead, they find that their simulations with momentum-driven winds (Murray et al. 2005; Zhang & Thompson 2012) best reproduce the slope, shape, scatter, and evolution of the MZR because the wind velocity scales with the escape velocity of the halo. Their model naturally produces a FMR that shows little evolution since $z = 3$, consistent with observations (Mannucci et al. 2010; Richard et al. 2011; Cresci et al. 2012). However, their FMR
does not quite reach the low observed scatter reported by Mannucci et al. (2010). Additionally, they find that the coefficient relating $M_\star$ and SFR that minimizes the scatter in the FMR is different from the one found by Mannucci et al. (2010). While there is hardly a consensus among galaxy evolution models about how to produce the MZR and $M_\star-Z-SFR$ relation, it is clear that additional observational constraints would improve the situation. So far, the overall normalization of the MZR and the $M_\star-Z-SFR$ relation have been mostly ignored by galaxy evolution models due to uncertainties in the nucleosynthetic yields used by the models and the large (up to a factor of five) uncertainties in the normalization of the observed relations caused by systematic offsets among metallicity calibrations. If these uncertainties could be reduced, then the normalization could be used as an additional constraint on galaxy evolution models.

The current metallicity and the metal enrichment history also have implications for certain types of stellar explosions. There is mounting evidence that long duration gamma ray bursts (Stanek et al. 2006), over-luminous type II supernovae (Stoll et al. 2011), and super-Chandrasekhar type Ia supernovae (Khan et al. 2011) preferentially occur in low metallicity environments. The progenitors of long gamma ray bursts and over-luminous type II supernovae are thought to be massive stars and the nature of their explosive death could plausibly depend on their metallicity. The cause of the association between super-Chandrasekhar type Ia supernovae and low metallicity environments is still highly uncertain because the progenitors are not well known.
Nevertheless, accurate absolute metallicities for the host galaxies of the progenitors of gamma ray bursts, over-luminous supernovae, and super-Chandrasekhar type Ia supernovae will help inform the models of stellar evolution and explosions that attempt to explain these phenomena.

The uncertainty in the absolute metallicity scale can be traced to differences between the two main methods of measuring metallicity: the direct method and strong line method. The direct method utilizes the flux ratio of auroral to strong lines to measure the electron temperature of the gas, which is a good proxy for metallicity because metals are the primary coolants of H II regions. This flux ratio is sensitive to temperature because the auroral and strong lines originate from the second and first excited states, respectively, and the relative level populations depend heavily on electron temperature. The electron temperature is a strong function of metallicity, such that hotter electron temperatures correspond to lower metallicities. In the direct method, the electron temperature estimate is the critical step because the uncertainty in metallicity is nearly always dominated by the uncertainty in the electron temperature. The strong line method uses the flux ratios of the strong lines, which do not directly measure the metallicity of the H II regions but are metallicity-sensitive and can be calibrated to give approximate metallicities. The direct method is chosen over strong line methods when the auroral lines can be detected, but these lines are often too weak to detect at high metallicity. The strong lines, on the other hand, are much more easily detected than the auroral lines,
particularly in metal-rich objects. Consequently, the strong line method can be used across a wide range of metallicity and on much lower signal-to-noise ratio (SNR) data, so nearly all metallicity studies of large galaxy samples employ the strong line method. Despite the convenience of the strong line method, the relationship between strong line ratios and metallicity is complicated due to the sensitivity of the strong lines to the hardness of the incident stellar radiation field and the excitation and ionization states of the gas. Thus, strong line ratios must be calibrated (1) empirically with direct method metallicities, (2) theoretically with photoionization models, or (3) semi-empirically with a combination of direct method metallicities and theoretically calibrated metallicities. Unfortunately, the three classes of calibrations do not generically produce consistent metallicities. For example, metallicities determined with theoretical strong line calibrations are systematically higher than those from the direct method or empirical strong line calibrations by up to \( \sim 0.7 \) dex (for a detailed discussion see Moustakas et al. 2010; Stasińska 2010). The various strong line methods also exhibit systematic disagreements as a function of metallicity and perform better or poorer in certain metallicity ranges.

The cause of the discrepancy between direct method metallicities and theoretically calibrated metallicities is currently unknown. As recognized by Peimbert (1967), the electron temperatures determined in the direct method might be overestimated in the presence of temperature gradients and/or fluctuations in \( \text{H} \ II \) regions. Such an effect would cause the direct method metallicities to be biased
low (Stasińska 2005; Bresolin 2008). A similar result could arise if the traditionally adopted electron energy distribution is different from the true distribution, as suggested by Nicholls et al. (2012). Alternatively, the photoionization models that serve as the basis for the theoretical strong line calibrations, such as CLOUDY (Ferland et al. 1998) and MAPPINGS (Sutherland & Dopita 1993), make simplifying assumptions in their treatment of H II regions that may result in overestimated metallicities, such as the geometry of the nebula or the age of the ionizing stars (see Moustakas et al. 2010, for a thorough discussion of these issues); however, no one particular assumption has been conclusively identified to be the root cause of the metallicity discrepancy.

In this work, we address the uncertainty in the absolute metallicity scale by using the direct method on a large sample of galaxies that span a wide range of metallicity. The uniform application of the direct method also provides more consistent metallicity estimates over a broad range in stellar mass. While the auroral lines used in the direct method are undetected in most galaxies, we have stacked the spectra of many galaxies (typically hundreds to thousands) to significantly enhance the SNR of these lines. In Section 2.2 we describe the sample selection, stacking procedure, and stellar continuum subtraction. Section 2.3 describes the direct method and strong line metallicity calibrations that we use. In Section 2.4 we demonstrate that mean galaxy properties can be recovered from stacked spectra. We show the electron temperature relations for the stacks in Section 2.3.1 and argue that
$T_\alpha[O\ II]$ is a better tracer of oxygen abundance than $T_\alpha[O\ III]$ in Section 2.3.2. Section 2.5 shows the main results of this study: the MZR and $M_\star-Z-SFR$ relation with the direct method. In Section 2.6, we present the direct method N/O relative abundance as a function of O/H and stellar mass. Section 2.7 details the major uncertainties in metallicity measurements and the implications for the physical processes that govern the MZR and $M_\star-Z-SFR$ relation. Finally, we present a summary of our results in Section 2.8. For the purpose of discussing metallicities relative to the solar value, we adopt the solar oxygen abundance of $12 + \log(O/H) = 8.86$ from Delahaye & Pinsonneault (2006). Throughout this work, stellar masses and SFRs are in units of $M_\odot$ and $M_\odot\,yr^{-1}$, respectively. We assume a standard ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,km\,s^{-1}\,Mpc^{-1}$.

2.2. Method

2.2.1. Sample Selection

The observations for our galaxy sample come from the SDSS Data Release 7 (DR7; Abazajian et al. 2009), a survey that includes $\sim$930,000 galaxies (Strauss et al. 2002) in an area of 8423 square degrees. The parent sample for this study comes from the MPA-JHU catalog of 818,333 unique galaxies which have derived stellar masses (Kauffmann et al. 2003a), SFRs (Brinchmann et al. 2004; Salim et al. 2007), and metallicities (T04). We chose only galaxies with reliable redshifts ($\sigma_z < 0.001$).

1Available at http://www.mpa-garching.mpg.de/SDSS/DR7/
in the range $0.027 < z < 0.25$ to ensure that the [O II] $\lambda$3727 line and the [O II] $\lambda\lambda$7320, 7330 lines fall within the wavelength range of the SDSS spectrograph (3800–9200 Å).

We discard galaxies classified as AGN because AGN emission line ratios may produce spurious metallicity measurements. We adopt the Kauffmann et al. (2003b) criteria (their Equation 1) to differentiate between star-forming galaxies and AGN, which employs the emission line ratios that define the Baldwin, Phillips, and Terlevich (1981) (BPT) diagram:

$$\log([\text{O III}]\lambda5007/H\beta) > 0.61[\log([\text{N II}]\lambda6583/H\alpha) - 0.05]^{-1} + 1.3.$$  (2.2)

We follow the T04 SNR thresholds for emission lines. Specifically, we restrict our sample to galaxies with $H\beta$, $H\alpha$, and [N II] $\lambda6583$ detected at $> 5\sigma$. Further, we apply the AGN–star-forming galaxy cut (Equation 2.2) to galaxies with $> 3\sigma$ detections of [O III] $\lambda5007$. We also select galaxies with [O III] $\lambda5007 < 3\sigma$ but $\log([\text{N II}]\lambda6583/H\alpha) < -0.4$ as star-forming to include high metallicity galaxies with weak [O III] $\lambda5007$.

At the lowest stellar masses ($\log[M_\ast] < 8.6$), this initial sample is significantly contaminated by spurious galaxies, which are actually the outskirts of more massive galaxies and were targeted due to poor photometric deblending. We remove galaxies whose photometric flags include DEBLEND_NOPEAK or DEBLENDED_AT_EDGE.
also visually inspected all galaxies with log($M_\star$) < 8.6 and discarded any that suffered from obvious errors in the stellar mass determination (again, likely as a result of off-center targeting of a much more massive galaxy).

After all of our cuts, the total number of galaxies in our sample is 208,529 and the median redshift is $z = 0.078$. At this redshift, the 3” diameter SDSS aperture will capture light from the inner 2.21 kpc of a galaxy. Since the central regions of galaxies will tend to be more metal-rich (Searle 1971), the metallicities measured from these observations will likely be biased high due to the aperture size relative to angular extent of the galaxies. However, we expect this bias is small for most galaxies (for a more detailed discussion see Tremonti et al. 2004; Kewley et al. 2005). In particular, the galaxies with very low stellar masses and metallicities that define the low mass end of the MZR tend to be compact and have homogeneous metallicities (e.g., Kobulnicky & Skillman 1997), although many of these are excluded by the criteria proposed by Kewley et al. (2005).

2.2.2. Stacking Procedure

The primary motivation for this investigation is to measure the metallicity of galaxies with the direct method. The main challenge is that the weak [O III] $\lambda 4363$ and [O II] $\lambda\lambda 7320, 7330$ auroral lines are undetected in most of the individual spectra. To improve the SNR of the spectra, we stacked galaxies that are expected to have similar metallicities and hence line ratios. Given the tightness of the MZR and
$M_\star-Z-SFR$ relation, it is reasonable to expect that galaxies at a given stellar mass, or simultaneously a given stellar mass and SFR, will have approximately the same metallicity. Thus, we have created two sets of galaxy stacks: (1) galaxies binned in 0.1 dex in $M_\star$ from log($M_\star/M_\odot$) = 7.0 to 11.0 (hereafter $M_\star$ stacks) and (2) galaxies binned in 0.1 dex in $M_\star$ from log($M_\star/M_\odot$) = 7.0 to 11.0 and 0.5 dex in SFR from log(SFR/[$M_\odot$ yr$^{-1}$]) = -2.0 to 2.0 (hereafter $M_\star-SFR$ stacks). We adopt the total stellar mass (Kauffmann et al. 2003a) and the total SFR (Brinchmann et al. 2004; Salim et al. 2007) values from the MPA-JHU catalog, as opposed to these quantities calculated only for the light within the fiber. For convenience, we will refer to the stacks by the type of stack with a subscript and a superscript to denote the upper and lower bounds of log($M_\star$) or log(SFR) (e.g., $M_\star^{8.7}$ is the $M_\star$ stack with log[$M_\star/M_\odot$] = 8.7–8.8, and SFR$^{0.5}$ corresponds to the $M_\star-SFR$ stacks with log[SFR/$M_\odot$ yr$^{-1}$] = 0.0–0.5). Figure 2.1 shows the number of galaxies in each $M_\star-SFR$ stack (each box represents a stack) with a measured metallicity (indicated by the color coding).

We stacked galaxy spectra that have been processed with the SDSS reduction pipeline (Stoughton et al. 2002). First, we corrected for Milky Way reddening with the extinction values from Schlegel et al. (1998). Then, the individual galaxy spectra were shifted to the rest frame with the redshifts from the MPA/JHU catalog. Next, we linearly interpolated the spectra onto a universal grid (3700–7360 Å; $\Delta \lambda = 1$ Å) in linear-$\lambda$ space. This interpolation scheme conserves flux in part because the
wavelength spacing of the grid is narrower than the width of bright emission lines. The spectra were then normalized to the mean flux from 4400–4450 Å. Finally, the spectra were co-added (i.e., we took the mean flux in each wavelength bin) to form the stacked spectra (see Section 2.4 for comparisons between the electron temperatures and metallicities of stacks and individual galaxies).

Figure 2.2 shows the SNR increase of the [O iii] λ4363 (left column), [N ii] λ5755 (middle column), and [O ii] λλ7320, 7330 (right column) lines as the spectra are processed from a typical single galaxy spectrum (top row) to the stacked spectrum (second row) to the stellar continuum subtracted spectrum (third row; see Section 2.2.3) or the narrow wavelength window stellar continuum subtracted spectrum (bottom row; see Section 2.2.3). The spectra in the top row are from a typical galaxy in the log($M_*$) = 8.7–8.8 bin; the bottom three rows show the stacked spectra from the same bin. In each panel, we report the continuum root mean square (rms). The decrease in the continuum noise when comparing the spectra in the top row to the second row of Figure 2.2 is dramatic. Further significant noise reduction can be achieved by removing the stellar continuum (shown in the bottom two rows of Figure 2.2), as we describe in Section 2.2.3

2.2.3. Stellar Continuum Subtraction

Stacking the spectra increases the SNR, but it is important to fit and subtract the stellar continuum to detect and accurately measure the flux of these lines,
especially $[\text{O } \text{iii}] \lambda 4363$ due to its proximity to the $H\gamma$ stellar absorption feature.

We subtracted the stellar continuum with synthetic template galaxy spectra created with the STARLIGHT stellar synthesis code (Cid Fernandes et al. 2005), adopted the Cardelli et al. (1989) extinction law, and masked out the locations of the emission lines. The synthetic spectra were created from a library of 300 empirical MILES spectral templates (Sánchez-Blázquez et al. 2006; Cenarro et al. 2007; Vazdekis et al. 2010; Falcón-Barroso et al. 2011, data as obtained from the MILES website\(^2\)). The MILES templates provided an excellent fit to the stellar continuum (see bottom two rows of Figure 2.2). We note the MILES templates yielded better fits to the very high SNR spectra than the Bruzual & Charlot (2003) spectral templates, based on the STELIB (Le Borgne et al. 2003) library.

We performed stellar template fits to the entire spectral range, select subregions centered on weak lines of interest, and subregions around the strong lines blueward of 4000 Å. The latter are situated among a forest of stellar absorption lines. The line fluxes of the strong emission lines redward of 4000 Å ($H\beta$, $[\text{O } \text{iii}] \lambda \lambda 4959, 5007$, $H\alpha$, $[\text{N } \text{ii}] \lambda \lambda 6548, 6583$, and $[\text{S } \text{ii}] \lambda \lambda 6716, 6731$) were measured from the spectrum where the stellar continuum was fit over the full wavelength range of our stacked spectra ($\lambda = 3600–7360$ Å; see third row of Figure 2.2). The stellar continuum subtraction near weak emission lines ($[\text{S } \text{ii}] \lambda 4069$, $[\text{O } \text{iii}] \lambda 4363$, He $\Pi \lambda 4686$, $[\text{N } \text{ii}] \lambda 5755$, $[\text{S } \text{iii}] \lambda 6312$, $[\text{Ar } \text{iv}] \lambda 4740$, and $[\text{O } \text{ii}] \lambda \lambda 7320, 7330$) and blue strong

\(^2\)http://miles.iac.es/
emission lines ([O II] $\lambda$3727 and [Ne III] $\lambda$3868) was improved if the stellar continuum fit was restricted to limited wavelength ranges within a few 100 Å of the line of interest (compare the third and bottom rows of Figure 2.2). For the weak lines and blue strong lines, we measured the line fluxes from the stellar continuum subtracted spectra within these narrow wavelength windows (details are listed in Table 1 of Andrews & Martini 2013). In order to compare the line fluxes across regions with different stellar continuum subtraction (e.g., from portions of the spectrum that were fit with smaller wavelength ranges), we denormalized the spectra after the starlight fit.

2.2.4. Automated Line Flux Measurements

We used the specfit task (Kriss 1994) in the IRAF/STSDAS package to automatically fit emission lines with a $\chi^2$ minimization algorithm. We simultaneously fit a flat continuum and Gaussian line profiles for the emission lines, even if lines were blended. For doublets, we fixed the width of the weaker line by pinning its velocity width to the stronger line ([O II] $\lambda$3726 to [O II] $\lambda$3729, [O III] $\lambda$4959 to [O III] $\lambda$5007, [N II] $\lambda$6548 to [N II] $\lambda$6583, [S II] $\lambda$6731 to [S II] $\lambda$6716, and [O II] $\lambda$7330 to [O II] $\lambda$7320). We also included the continuum rms of the spectrum as an input to the fitting procedure. After experimenting with several different $\chi^2$ minimization algorithms implemented within specfit, we chose the simplex algorithm because of its consistent convergence, particularly for weak lines. Line fluxes measured by specfit
generally agreed well with line fluxes measured interactively with the OSU LINER package. The uncertainty in the line flux is based on the $\chi^2$ fit returned from specfit.

Finally, all line fluxes were corrected for reddening with the extinction law from Cardelli et al. (1989) and the assumption that the intrinsic ratio of the Balmer lines is set by case B recombination ($H\alpha/H\beta = 2.86$ for $T_e = 10,000$ K). We adopted a fixed $H\alpha/H\beta$ ratio, even though it is a weak function of electron temperature. For the log($M*/M_\odot$) = 10.0–10.1 stack ($T_e [O II] = 7200$ K), whose oxygen abundance is dominated by O$^+$ (i.e., a stack where the potential effect would be maximal due to the long wavelength baseline between [O II] $\lambda$3727 and [O II] $\lambda\lambda$7320, 7330), this effect would decrease log(O$^+$/H$^+$) by $\sim$0.07 dex. The line fluxes are presented in an online table whose columns are described in Table 2 of Andrews & Martini 2013.

We disregarded lines that were poorly fit (negative flux, uncertainty in central wavelength >1 Å, had uncertainty in the velocity width of >100 km/s, or had low SNR [< 5$\sigma$]). Further care was taken to ensure the robustness of [O III] $\lambda$4363 flux measurements. As $M_*$ increased to moderate values (log($M_*$) > 9.0), an unidentified emission feature at 4359 Å became blended with the [O III] $\lambda$4363 line, which limited the SNR of the line flux measurement independent of the continuum rms. We are unsure of the origin of this feature, but it could be caused by an over-subtraction in the stellar continuum fit. We simultaneously fit the 4359 Å feature and [O III] $\lambda$4363 and pinned the velocity width of both lines to $H\gamma$. If 4359Å > 0.5 [O III] $\lambda$4363, then we determined that [O III] $\lambda$4363 could not be robustly fit. If [O III] $\lambda$4363...
could be well fit, we refit it with a single Gaussian whose velocity width was pinned to H\textsubscript{\gamma}. The line flux measurements from the single Gaussian fitting agreed better with interactive line flux measurements than the deblended line flux measurements. The remaining weak lines are in regions without strong stellar absorption features. Often, the [O \textsc{ii}] λ\lambda 7320, 7330 lines could be detected in the stacked spectra without the stellar continuum fit (see Figure 2.2f). The [N \textsc{ii}] λ5755 and [S \textsc{ii}] λ4069 auroral lines were usually too weak to be detected without stellar continuum subtraction.

Optical recombination lines, such as C \textsc{ii} λ4267 and O \textsc{ii} λ4649, are also sensitive to metallicity. Unlike auroral lines, they are almost independent of temperature, so they could provide a useful check on the direct method metallicities. Unfortunately, optical recombination lines tend to be very weak (e.g., the median O \textsc{ii} λ4649/[O \textsc{iii}] λ4363 ratio of five extragalactic H \textsc{ii} regions studied by Esteban et al. 2009 was 0.08), and we did not detect them in the stacked spectra.

2.3. Electron Temperature and Direct Abundance Determination

2.3.1. Electron Temperatures

Different ionic species probe the temperature of different ionization zones of H \textsc{ii} regions (e.g., Stasińska 1982; Garnett 1992). In the two-zone model, the high ionization zone is traced by [O \textsc{iii}], and the low ionization zone is traced by [O \textsc{ii}],
Campbell et al. (1986) used the photoionization models of Stasińska (1982) to derive a linear relation between the temperatures in these zones,

\[ T_e[\text{O ii}] = T_e[\text{N ii}] = T_e[\text{S ii}] = 0.7T_e[\text{O iii}] + 3000, \]

where \( T_e \) is in units of K. Subsequently, we will refer to this relation as the \( T_2-T_3 \) relation (see Pagel et al. 1992 and Izotov et al. 2006 for alternative formulations of the \( T_2-T_3 \) relation). This relation is especially useful to infer the abundance of unseen ionization states, a critical step in measuring the total oxygen abundance.

While convenient, this theoretical relation may be one of the biggest uncertainties in the direct method because it is not definitively constrained by observations due to the large random errors in the flux of \([\text{O II}] \lambda\lambda 7320, 7330\) (e.g., see Kennicutt et al. 2003; Pilyugin et al. 2006). The high SNR of our stacked spectra enables us to measure the electron temperature of both the high and low ionization zones for many of our stacks.

We measured the electron temperature of \([\text{O III}], [\text{O II}], [\text{N II}], \) and \([\text{S II}] \) with the \textit{nebular.temden} routine (Shaw & Dufour 1995) in IRAF/STSDAS, which is based on the five level atom program of De Robertis et al. (1987). This routine determines the electron temperature from the flux ratio of the auroral to strong emission line(s) for an assumed electron density. The diversity of these temperature diagnostics are valuable cross-checks and provide an independent check on the applicability of the \( T_2-T_3 \) relation; however, for measuring oxygen abundances, we only use \( T_e[\text{O III}] \) and
$T_e[\text{O}\ II]$. The electron density ($n_e$) can be measured from the density sensitive $[\text{S}\ II]$ $\lambda\lambda 6716, 6731$ doublet (cf., Cai & Pradhan 1993). For 6/45 of the $M_*$ stacks and 65/228 of the $M_*-$SFR stacks, $[\text{S}\ II] \lambda 6716 / [\text{S}\ II] \lambda 6731$ was above the theoretical maximum ratio of 1.43 (Osterbrock 1989), which firmly places these galaxies in the low density regime, and we assume $n_e = 100$ cm$^{-3}$ for our analysis. Yin et al. (2007) found similar inconsistencies between the theoretical maximum and measured flux ratios for individual galaxies, which suggests that there might be a real discrepancy between the maximum observed and theoretical values of $[\text{S}\ II] \lambda 6716 / [\text{S}\ II] \lambda 6731$.

We calculated the electron temperature and density uncertainties by propagating the line flux uncertainties with Monte Carlo simulations. For the simulations, we generated 1,000 realizations of the line fluxes (Gaussian distributed according to the 1σ uncertainty) and processed these realizations through `nebular.temden`. The electron temperatures of the stacks are given in Table 3 of Andrews & Martini (2013) (full version available online).

In Figure 2.3 we plot the electron temperatures of $[\text{O}\ II]$, $[\text{N}\ II]$, and $[\text{S}\ II]$ against the $[\text{O}\ III]$ electron temperature for the $M_*$ stacks (left column; open circles) and the $M_*-$SFR stacks (right column; circles color-coded by SFR). For comparison, we show the $T_2-T_3$ relation (Equation 2.3) as the black line in each panel. In all three $T_e-T_e$ plots, the $M_*$ stacks form a tight locus that falls within the distribution of $M_*-$SFR stacks. The $M_*-$SFR stacks show a large dispersion in $T_e[\text{O}\ II]$ at fixed $T_e[\text{O}\ III]$ that is not present in the $M_*$ stacks. Most of this scatter is due to stacks
with SFR$^{1.5}_{1.0}$, which approach and exceed the $T_e[\text{O II}]-T_e[\text{O III}]$ relation. On the other hand, the $M_\star$–SFR stacks show little scatter in the $T_e[\text{N II}]-T_e[\text{O III}]$ and $T_e[\text{S II}]-T_e[\text{O III}]$ plots, and they track the $M_\star$ stacks in these plots.

The vast majority of the stacks in Figure 2.3 fall below the $T_2-T_3$ relation, independent of the type of stacks ($M_\star$ or $M_\star$–SFR) or the tracer ion ([O II], [N II], or [S II]). The multiple temperature indicators show that the $T_2-T_3$ relation overpredicts the temperature in the low ionization zone (or underpredicts the temperature in the high ionization zone). If we assume that $T_e[\text{O III}]$ is accurate (i.e., the temperature in the low ionization zone is overestimated by the $T_2-T_3$ relation), then the median offsets from the $T_2-T_3$ relation for the $M_\star$ stacks and the $M_\star$–SFR stacks, respectively, are

- $T_e[\text{O II}]$: $-2000$ K and $-1300$ K,
- $T_e[\text{N II}]$: $-1200$ K and $-1400$ K,
- $T_e[\text{S II}]$: $-4100$ K and $-3300$ K.

The $T_e[\text{O II}]$ and $T_e[\text{N II}]$ offsets from the $T_2-T_3$ relation for the $M_\star$ stacks are consistent given the scatter, which suggests that the $T_2-T_3$ relation overestimates the low ionization zone $T_e$ by $\sim 1000–2000$ K. The $T_e[\text{S II}]$ measurements show a larger offset from the $T_2-T_3$ relation than $T_e[\text{O II}]$ and $T_e[\text{N II}]$. The outlier in the $M_\star$–SFR $T_e[\text{S II}]-T_e[\text{O III}]$ panel also has a high $T_e[\text{O II}]$, but this outlier just corresponds to
a single galaxy, so it may not be representative of all galaxies with this stellar mass and SFR.

The offset between the electron temperatures of the stacks and the $T_2-T_3$ relation is analogous to the trend for individual galaxies found by Pilyugin et al. (2010), which persists when these galaxies are stacked (see Section 2.4 and Figure 2.8). The similar distributions of stacks and individual galaxies relative to the $T_2-T_3$ relation shows that the offset for the stacks is not a by-product of stacking but rather a reflection of the properties of the individual galaxies (for further discussion see Section 2.4).

At high SFRs ($\text{SFR}_{1.5}$ and $\text{SFR}_{2.0}$), the offset in $T_e[\ion{O}{II}]$ disappears, and the median $T_e[\ion{O}{II}]$ of these stacks is consistent with the $T_2-T_3$ relation, albeit with a large dispersion. The emission from these galaxies is likely dominated by young stellar populations, whose hard ionizing spectrum may be similar to the single stellar spectra used by Stasińska (1982) to model H II regions. However, a single stellar effective temperature may not be appropriate for galaxy spectra that include a substantial flux contribution from older H II regions that have softer ionizing spectra (Kennicutt et al. 2000; Pilyugin et al. 2010).

Figure 2.4 compares the electron temperatures of $[\ion{O}{II}]$ and $[\ion{N}{II}]$ for the $M_\star$-SFR stacks (color-coded by SFR). In the two-zone model, both $T_e[\ion{O}{II}]$ and $T_e[\ion{N}{II}]$ represent the temperature of the low ionization zone, so these temperatures
should be the same. The stacks scatter around the line of equality (black line),
though the median offset from the \( T_{e}[\text{O II}] = T_{e}[\text{N II}] \) relation is 1100 K towards
higher \( T_{e}[\text{N II}] \). If only the stacks that also have detectable \( [\text{O III}] \lambda 4363 \) are
considered (most of which have \( T_{e}[\text{O II}] \gtrsim 8000 \) K), then the median offset from the
relation is smaller than the median uncertainty on \( T_{e}[\text{N II}] \). The agreement between
\( T_{e}[\text{O II}] \) and \( T_{e}[\text{N II}] \) for this subset of stacks is consistent with the similar offsets
found for \( T_{e}[\text{O II}] \) and \( T_{e}[\text{N II}] \) relative to the \( T_{2} - T_{3} \) relation in Figure 2.3.

2.3.2. Ionic and Total Abundances

We calculated the ionic abundance of O\(^+\) and O\(^{++}\) with the \textit{nebular.ionic} routine
(De Robertis et al. 1987; Shaw & Dufour 1995) in \textsc{iraf/stsdas}, which determines
the ionic abundance from the electron temperature, electron density, and the flux
ratio of the strong emission line(s) relative to H\(\beta\). We derived the ionic abundance
uncertainties with the same Monte Carlo simulations used to compute the electron
temperature and density uncertainties (see Section 2.3.1); the ionic abundance
uncertainties were propagated analytically to calculate the total abundance
uncertainties. We do not attempt to correct for systematic uncertainties in the
absolute abundance scale.

The top two panels of Figure 2.5 show the ionic abundance of O\(^+\) and O\(^{++}\) as
a function of stellar mass for the \( M_{*} \) stacks (open circles) and the \( M_{*} - \text{SFR} \) stacks
(circles color-coded by SFR). The O\(^+\) abundance increases with stellar mass at
fixed SFR and decreases with SFR at fixed stellar mass. The abundance of O$^{++}$ is relatively constant as a function of stellar mass but is detected in galaxies with progressively higher SFRs as stellar mass increases.

In Figure 2.5c, we plot the logarithmic ratio of the O$^{++}$ and O$^+$ abundances as a function of stellar mass. The dotted line in Figure 2.5c shows equal abundances of O$^+$ and O$^{++}$. The contribution of O$^+$ to the total oxygen abundance increases with stellar mass at fixed SFR and decreases with SFR at fixed stellar mass (i.e., in the same sense as how the O$^+$ abundance changes with $M_\star$ and SFR). The O$^+$ abundance dominates the total oxygen abundance in the majority of the stacks (i.e., above log[$M_\star$] = 8.2 for the $M_\star$ stacks and in half of the $M_\star$–SFR stacks with detected [O II] $\lambda\lambda7320, 7330$). Furthermore, the O$^+$ abundance can be measured in many high stellar mass and/or low SFR stacks that lack a measured O$^{++}$ abundance, which indicates that O$^+$ is very likely the main ionic species of oxygen in these stacks too. A simple extrapolation of the log(O$^{++}$/O$^+$) ratio to higher stellar masses for the $M_\star$ stacks shows that the O$^{++}$ abundance would contribute less than 10% of the total oxygen abundance.

We assume that the total oxygen abundance is the sum of the ionic abundances of the two dominant species,

$$\frac{O}{H} = \frac{O^+}{H^+} + \frac{O^{++}}{H^+},$$

(2.4)
and the total abundance uncertainties were determined by propagating the ionic abundance uncertainties. In highly ionized gas, oxygen may be found as O$^{3+}$, but its contribution to the total oxygen abundance is minimal. Abundance studies that use the direct method typically measure $T_e[\text{O III}]$ and the O$^{++}$ abundance but adopt the $T_2-T_3$ relation to infer $T_e[\text{O II}]$ and the O$^+$ abundance. However, Figure 2.3 shows that the $T_2-T_3$ relation overestimates $T_e[\text{O II}]$, which leads to an underestimate of the O$^+$ abundance and the total oxygen abundance. Many of the stacks have measured O$^+$ and O$^{++}$ abundances, so the total oxygen abundance can be measured accurately in these stacks without using the $T_2-T_3$ relation.

To extend our total oxygen abundance measurements to higher stellar mass, we form a “composite” metallicity calibration (see Figure 2.6) that uses the O$^+$ and O$^{++}$ abundances when available and the O$^+$ abundance plus the O$^{++}$ abundance inferred with the $T_2-T_3$ relation if $T_e[\text{O II}]$ is measured but not $T_e[\text{O III}]$ (in the opposite sense from how it is normally applied). The total oxygen abundance of the latter group of stacks is dominated by the O$^+$ abundance, so the inferred O$^{++}$ abundance makes only a small contribution ($<10\%$ based on the trend indicated by Figure 2.5c). A simple combination of these two metallicity calibrations would lead to a discontinuity at their interface (in the MZR) because applying the $T_2-T_3$ relation underestimates $T_e[\text{O III}]$ and thus overestimates the O$^{++}$ abundance. To account for this effect, we decrease the total oxygen abundances that adopt the $T_2-T_3$ relation by the median offset between the two calibrations where they are both
measured (0.18 dex for the $M_\star$ stacks). For the $M_\star$–SFR stacks, we calculate the median offset for each SFR bin (reported in Table 4 of Andrews & Martini 2013). The offsets are nearly constant as a function of $M_\star$ and stem from the approximately constant offset in the $T_e[O\,\text{II}]$–$T_e[O\,\text{III}]$ plot (top row of Figure 2.3). Because we account for the systematic offset from the $T_2$–$T_3$ relation, our composite metallicities are insensitive to the exact choice of the $T_2$–$T_3$ relation. The metallicities of the stacks are presented in Table 3 of Andrews & Martini (2013).

Most direct method metallicity studies measure the $[O\,\text{III}]$ λ4363 line flux but not the $[O\,\text{II}]$ λλ7320, 7330 line fluxes, so they must adopt a $T_e[O\,\text{II}]$–$T_e[O\,\text{III}]$ relation, such as the $T_2$–$T_3$ relation, to estimate the O$^+$ abundance. One reason for this is the large wavelength separation between the $[O\,\text{II}]$ λ3727 strong line and the $[O\,\text{II}]$ λλ7320, 7330 auroral lines used to measure $T_e[O\,\text{II}]$. The flux ratio of these two line complexes can be affected by a poor reddening correction, particularly for the $[O\,\text{II}]$ λ3727 line, and some spectrographs cannot observe this entire wavelength range efficiently. In individual spectra, the $[O\,\text{II}]$ λλ7320, 7330 lines can be overwhelmed by the noise, which can lead to a large scatter in the $T_e[O\,\text{II}]$–$T_e[O\,\text{III}]$ diagram (see Figure 1 of Kennicutt et al. 2003 or Figure 4 of Izotov et al. 2006). Fortunately, the noise near $[O\,\text{II}]$ λλ7320, 7330 appears to be random and is effectively reduced by stacking, even without the stellar continuum subtraction (see Figure 2.2f). Kennicutt et al. (2003) noted that the $[O\,\text{II}]$ λλ7320, 7330 line fluxes may be affected by recombination of O$^{++}$, although they find that the typical contribution to the
[O II] λλ7320, 7330 line fluxes is <5% (based on the correction formulae from Liu et al. 2000) and that $T_e[O II]$ is affected by ~2–3%, which corresponds to <400 K for the H II regions in their study.

We also calculated the ionic abundance of N$^+$ with nebular.ionic, similar to the procedure used to calculate the ionic abundances of O$^+$ and O$^{++}$, except that we adopt $T_e[O II]$ as the electron temperature instead of $T_e[N II]$ because the [O II] λλ7320, 7330 lines are detected in more stacks and with higher SNR than the [N II] λ5755 line (see Figure 2.2). The relative ionic abundance of N$^+/O^+$ was derived from the ionic abundances of each species. We then assume that N/O = N$^+/O^+$ (Peimbert & Costero 1969; Garnett 1990) to facilitate comparison with other studies in the literature (e.g., Vila Costas & Edmunds 1993). Although this assumption is uncertain, Nava et al. (2006) found that it should be accurate to ~10% for low metallicity objects (12 + log[O/H] ≤ 8.1). The N/O abundances of the stacks are reported in Table 3 of Andrews & Martini (2013).

2.3.3. Strong Line Metallicities

We compare our direct method metallicities to strong line metallicities with various empirical and theoretical calibrations of the most common line ratios:

- $R_{23}$: ([O II] λ3727 + [O III] λλ4959, 5007) / Hβ

- N2O2: [N II] λ6583 / [O II] λ3727,
• N2: [N II] λ6583 / Hα,


We derived metallicities for our stacks with the theoretical $R_{23}$ calibrations of McGaugh (1991, hereafter M91), Zaritsky et al. (1994, hereafter Z94), and Kobulnicky & Kewley (2004, hereafter KK04); the hybrid empirical–theoretical N2 calibration of Denicoló et al. (2002, hereafter D02); the theoretical N2O2 calibration of Kewley & Dopita (2002, hereafter KD02); and the mostly empirical N2 and O3N2 calibrations of Pettini & Pagel (2004, hereafter PP04). We determined uncertainties on the strong line metallicities with the Monte Carlo simulations detailed in Section 2.3.1; these uncertainties do not account for systematic uncertainties in the absolute abundance scale. For a detailed discussion of these calibrations and formulae to convert between the metallicities derived from each calibration see Kewley & Ellison (2008).

2.4. How Does Stacking Affect Measured Electron Temperatures and Metallicities?

Stacking greatly increases SNR and thus enables measurements of physical properties that are unattainable for individual objects. However, measurements from stacked spectra are only meaningful if they represent the typical properties of the objects that went into the stack. To evaluate the effect of stacking on the electron temperatures
and metallicities, we stacked a sample of 181 SDSS DR6 (Adelman-McCarthy et al. 2008) galaxies with individual detections of [O \textsc{iii}] $\lambda$4363 and [O \textsc{ii}] $\lambda\lambda$7320, 7330 from Pilyugin et al. (2010) in bins of 0.1 dex in stellar mass. Figure 2.7 shows the $T_e$[O \textsc{iii}], $T_e$[O \textsc{ii}], and the direct method metallicities of the individual galaxies (gray squares) and stacks (black circles) relative to the mean of the galaxies that went into each stack. For all three properties, the stacks are consistent with the mean of the galaxies within the measurement uncertainties, which demonstrates that the properties derived from galaxies stacked in narrow bins of stellar mass are representative of the mean properties of the input galaxies.

In Figure 2.8, the [O \textsc{ii}], [N \textsc{ii}], and [S \textsc{ii}] electron temperatures are plotted as a function of the [O \textsc{iii}] electron temperature for the galaxies (squares color-coded by SFR) and stacks (black circles). The black line in each panel indicates the $T_2$–$T_3$ relation (Equation 2.3). The stacks fall within the distribution of galaxies in the $T_e$[O \textsc{ii}]–$T_e$[O \textsc{iii}] and $T_e$[S \textsc{ii}]–$T_e$[O \textsc{iii}] plots (Figure 2.8a,c). There is some discrepancy between the stacks and galaxies in the $T_e$[N \textsc{ii}]–$T_e$[O \textsc{iii}] plot (Figure 2.8b), but the paucity of [N \textsc{ii}] $\lambda$5755 detections limits the usefulness of any strong conclusions based on $T_e$[N \textsc{ii}]. Overall, the qualitative agreement between the electron temperatures of the stacks and galaxies, especially for $T_e$[O \textsc{ii}] and $T_e$[S \textsc{ii}], demonstrates that the offset from the $T_2$–$T_3$ relation for the stacks shown in Figure 2.3 is not an artifact of stacking.
The majority of the galaxies lie below the $T_2-T_3$ relation, as was previously shown by Pilyugin et al. (2010). We find a similar result for the galaxies in the $T_e$[S II]$-T_e$[O III] relation. Galaxies with moderate SFRs ($\log[\text{SFR}] \sim 0.0$) are preferentially further below the $T_2-T_3$ relation than galaxies with high SFRs ($\log[\text{SFR}] \gtrsim 1.0$) in the $T_e$[O II]$-T_e$[O III] plot. A similar effect is also present in the $M_*$–SFR stacks. Pilyugin et al. (2010) found that galaxies with lower excitation parameters and [O III] $\lambda 5007/$H$\beta$ flux ratios had larger offsets from the $T_2-T_3$ relation, which is consistent with our result based on SFR. They showed that the offset from the $T_2-T_3$ relation is likely due to the combined emission from multiple ionizing sources by comparing the observed $T_e$[O II]$-T_e$[O III] relation with the temperature predicted by H II region models that include ionizing sources of various temperatures. Based on these models, they concluded that differences in the hardness of the ionizing radiation, caused by the age-dependence of H II region spectral energy distributions, govern the scatter in the $T_e$[O II]$-T_e$[O III] plot for their sample of galaxies. Both our results and theirs suggest that galaxies with higher SFRs are more similar to the H II region models that served as the basis for the $T_2-T_3$ relation than galaxies with moderate SFRs. This is because they are more likely to be dominated by younger stellar populations that are better approximated by the input to the Stasińska (1982) models (see Section 2.3.1 for additional discussion).

The electron temperatures and metallicities of the stacks are unbiased relative to those of the input galaxies, but there is some evidence that the integrated
galaxy electron temperature and metallicity are systematically higher and lower, respectively, than the electron temperatures and metallicities of the individual H II regions in the galaxy. Kobulnicky et al. (1999) compared the electron temperatures and metallicities of individual H II regions in a galaxy to the pseudo-global values derived by stacking the spectra of the individual H II regions. They showed that the electron temperatures and direct method metallicities of their galaxies were biased towards higher temperatures and lower metallicities by $\sim 1000$–$3000$ K and $0.05$–$0.2$ dex, respectively, relative to the median values of the individual H II regions. Global spectra are biased because they are the luminosity-weighted average of the H II regions, whose properties can vary widely (see, e.g., the large scatter around the $T_2$–$T_3$ relation for $T_e$ measurements of individual H II regions in Figure 1 of Kennicutt et al. 2003 or Figure 4 of Izotov et al. 2006). The fluxes of the auroral lines might be particularly affected by a luminosity-weighted average because auroral line flux decreases non-linearly with metallicity. While Kobulnicky et al. (1999) only studied the effects on [O III] $\lambda 4363$, the relative contribution of each H II region likely varies among the commonly measured ionic species, potentially yielding results that do not agree with the $T_2$–$T_3$ relation. We also note that their method of stacking H II regions does not perfectly simulate global line flux measurements because it does not account for the contribution of diffuse ionized gas (i.e., the emission from gas not in H II regions), which may affect the [N II] and [S II] line fluxes (see Moustakas & Kennicutt 2006). In summary, the differences in electron temperatures and
metallicities between galaxy stacks and individual H II regions are dominated by the systematic offset between global galaxy properties and individual H II regions rather than any effects from stacking the global galaxy spectra.

The auroral lines are undetectable in high stellar mass galaxies, so we investigate the effect of stacking by comparing the oxygen strong line fluxes of individual galaxies to the stack of those galaxies. Figure 2.9 shows the [O II] λ3727 and [O III] λ5007 fluxes relative to Hβ for individual galaxies (small black and blue circles) with log(M*) = 10.5–10.6 and log(SFR) = 1.0–1.5 and the stack of the same galaxies (large green circle). The small black and blue circles correspond to the line fluxes determined with our pipeline and the MPA-JHU pipeline, respectively. The distribution of individual galaxies with fluxes measured by our pipeline and the MPA-JHU pipeline coincide well. In detail, the median fluxes from our pipeline are 0.08 and 0.04 dex higher for [O III] λ5007 and [O II] λ3727, respectively, than the median fluxes from the MPA-JHU pipeline. The [O III] λ5007 and [O II] λ3727 fluxes of the stack are 0.09 and 0.01 dex higher, respectively, than the median of fluxes from our pipeline. Although the spread is large in the individual galaxies (<1 dex for both [O III] λ5007 and [O II] λ3727), the stack is representative of the typical line fluxes of individual galaxies that went into the stack.

We also note that many of our stacks contain far more galaxies than are needed to simply detect a given line, and thus are unlikely to be dominated by a few, anomalous galaxies. As an example, we estimate how many galaxies would
need to be stacked for a detection of \([\text{O} \, \text{II}] \, \lambda\lambda 7320, \, 7330\). If we assume that the uncertainty on the line flux decreases as \(\sqrt{N_{\text{galaxies}}}\), the error on the measurement of any individual galaxy is \(\sigma_{\text{stack}} \ast \sqrt{N_{\text{galaxies}}}\). We use a 5\(\sigma\) detection threshold, so the minimum number of galaxies needed to detect a line is \(N = \left\lfloor \frac{(5\sigma)/\text{flux}}{2} \right\rfloor\). For the \(M_{\ast,0.5} \, \text{SFR}_{0.0}\) stack, the minimum number of galaxies required to detect \([\text{O} \, \text{II}] \, \lambda\lambda 7320, \, 7330\) is \(N_{\text{galaxies}} = 40\), which is well below the actual number of galaxies (1996) in this stack.

2.5. The Mass–Metallicity Relation and Mass–Metallicity–SFR Relation

2.5.1. The Mass–Metallicity Relation

In Figure 2.10, we plot the MZR with direct method metallicities for the \(M_{\ast}\) stacks (circles). We fit the MZR for the \(M_{\ast}\) stacks (black line) with the asymptotic logarithmic formula suggested by Moustakas et al. (2011):

\[
12 + \log(O/H) = 12 + \log(O/H)_{\text{asm}} - \log \left(1 + \left(\frac{M_{\text{TO}}}{M_{\ast}}\right)^{\gamma}\right),
\]

(2.5)

where \(12 + \log(O/H)_{\text{asm}}\) is the asymptotic metallicity, \(M_{\text{TO}}\) is the turnover mass, and \(\gamma\) controls the slope of the MZR. This functional form is preferable to a polynomial because polynomial fits can produce unphysical anticorrelations between mass and metallicity, particularly when extrapolated beyond the mass range over which they
were calibrated. The metallicities and fit parameters for the stacks are reported in Tables 3 and 4 of Andrews & Martini (2013), respectively. For comparison, we show the robust cubic polynomial fits of eight strong line MZRs (colored lines) from Kewley & Ellison (2008) in Figure 2.10a. The T04, Z94 R_{23}, KK04 R_{23}, KD02 N2O2, and M91 R_{23} MZRs are based on theoretical calibrations, whereas the D02 N2, PP04 O3N2, and PP04 N2 MZRs are based on empirical calibrations. In Figure 2.10b, the solid, dashed, and dotted gray lines indicate the median, 68% contour, and 95% contour, respectively, of the T04 MZR.

The most prominent aspect of the direct method MZR is its extensive dynamic range in both stellar mass and metallicity. It spans three decades in stellar mass and nearly one decade in metallicity; this wide range is critical for resolving the turnover in metallicity with a single diagnostic that is a monotonic relation between line strength and metallicity. The broad range in galaxy properties includes the turnover in the MZR, which is the first time this feature has been measured with metallicities derived from the direct method. Our stacked spectra also extend the direct method MZR to sufficiently high masses that there is substantial overlap with strong line measurements, and we use this overlap to compare them.

The direct method MZR shares some characteristics with strong line MZRs but differs in important ways, as can be seen in Figure 2.10a. The low mass end of the direct method MZR starts at log(M_*) = 7.4, a full decade lower than the strong line MZRs. Nonetheless, naive extrapolations of the T04, D02, PP04,
and PP04 MZRs are in reasonable agreement with our direct method MZR. At a stellar mass of \( \log(M_\star) = 8.5 \), the lowest stellar mass where strong line MZRs are reported, the direct method MZR is consistent with the T04 and the D02 MZRs. Above this mass, the direct method MZR and the D02 MZR diverge from the T04 MZR. At \( \log(M_\star) = 8.9 \), the direct method MZR turns over. By contrast, the strong line MZRs turns over at a much higher stellar mass (\( \log[M_\star] \sim 10.5 \)): a significant difference that has implications for how the MZR is understood in a physical context, which we discuss in Section 2.7.4. At high mass, the direct method MZR is in good agreement with the empirical strong line calibration MZRs, but the theoretical T04, Z94, KK04, and KD02 strong line calibration MZRs are offset to higher metallicities by \( \sim 0.3 \) dex at \( \log(M_\star) = 10.5 \), the highest mass stack with detected auroral lines. Figure 2.10b shows the direct method MZR in relation to the scatter of the T04 MZR. The direct method MZR is slightly below the median T04 MZR at \( \log(M_\star) = 8.5 \), crosses the 16th percentile at \( \log(M_\star) = 9.0 \), and drops below the 2nd percentile at \( \log(M_\star) = 9.9 \). Formally, the direct method MZR has a scatter of \( \sigma = 0.05 \) dex, but this value is not directly comparable to the scatter in the T04 MZR because stacking effectively averages over all galaxies in a bin, which erases information about galaxy-to-galaxy scatter.

At low masses (\( \log[M_\star] = 7.4-8.9 \); i.e., below the turnover), the direct method MZR scales as approximately \( O/H \propto M_\star^{1/2} \). While a comparison over the same mass range is not possible for the T04 MZR, its low mass slope, as determined from
\log(M_*) = 8.5-10.5, \text{ is shallower with } O/H \propto M_*^{1/3}. \text{ The discrepancy in the low mass slopes between the direct method and the T04 MZRs could be reasonably explained by the difference in the mass ranges over which the slopes were measured if the MZR steepens with decreasing stellar mass (cf., Lee et al. 2006). We note that the direct method and D02 MZRs have similar slopes and normalizations over a wide range in masses from } \log(M_*) = 8.5-10.0.

2.5.2. Mass–Metallicity–SFR Relation

The features of the direct method MZR are shaped by the SFR-dependence of the MZR, which we investigate with the $M_*-$SFR stacks. Figure 2.11 shows the $M_*-$SFR stacks (circles color-coded by SFR) in the mass–metallicity plane (see Figure 2.1 for the number of galaxies per stack). The solid colored lines indicate the asymptotic logarithmic fits (Equation 2.5) of the $M_*-$SFR stacks of a given SFR, hereafter referred to as SFR tracks (e.g., the orange line is the SFR$^{-0.5}$ track). The solid black line is the direct method MZR of the $M_*$ stacks from Figure 2.10; the solid, dashed, and dotted gray lines are the median, 68% contour, and 95% contour, respectively, of the T04 MZR. The error bars represent the mean error for the $M_*-$SFR stacks of a given SFR.

The $M_*-$SFR stacks help establish the robustness of the direct method MZR. The low turnover mass and metallicity of the direct method MZR relative to the T04 and other theoretical strong line calibration MZRs is reminiscent of empirical strong
line calibration MZRs that suffer from a lack of sensitivity at high metallicities. However, the most metal-rich $M_\star$–SFR stacks have some of the highest direct method metallicities ($12 + \log(O/H) > 9.0$)—metallicities well above the turnover metallicity of the direct method MZR. These measurements unambiguously demonstrate that the turnover in the direct method MZR is not caused by a lack of sensitivity to high metallicities.

The $M_\star$–SFR stacks also can be used to test if galaxies with the highest SFRs at a given stellar mass disproportionately influence the line fluxes and metallicities of the $M_\star$ stacks. High SFR galaxies have more luminous emission lines and lower metallicities and thus may dominate the inferred metallicity of the stack. To investigate this possibility, we calculated the difference between the metallicity of the $M_\star$ stack and the galaxy number-weighted average of the metallicities of the $M_\star$–SFR stacks (for the stacks with measured metallicities) at a given stellar mass. The median offset is only $-0.037$ dex in metallicity; for reference, the median metallicity uncertainties for the $M_\star$ and $M_\star$–SFR stacks are 0.019 and 0.027 dex, respectively. The slight offset could be due to preferentially including the metallicities of $M_\star$–SFR stacks with higher SFR (lower metallicity) relative to lower SFR (higher metallicity) in the weighted average because the former tend to have larger line fluxes than the latter, whereas the $M_\star$ stacks include the contribution from galaxies of all SFRs at a given stellar mass. Still, the magnitude of this offset is small, which indicates that the highest SFR galaxies do not have an appreciable effect on the metallicity.
of the \( M_* \) stacks because they are quite rare (see Figure 2.11). Furthermore, the metallicities of the \( M_* \) stacks effectively track the metallicity of the most common galaxies at a given stellar mass.

The most striking features of Figure 2.11 are the 0.3–0.6 dex offsets in metallicity at fixed stellar mass between the \( M_* \text{–SFR} \) stacks. This trend results from the substantial, nearly monotonic dependence of the MZR on SFR. At a given stellar mass, higher SFR stacks almost always have lower metallicities than lower SFR stacks, so there is little overlap between the different SFR tracks. Furthermore, the small regions with overlap may be the result of the observational uncertainties.

The interplay between stellar mass, SFR, and metallicity for typical galaxies is reflected in the features of the direct method MZR, especially the turnover mass. The constant SFR tracks (colored lines in Figure 2.11) show that metallicity increases with stellar mass at fixed SFR. However, the typical SFR also increases with stellar mass, which shifts the “typical” galaxy (as measured by the \( M_* \) stacks) to progressively higher SFR and consequently lower metallicity at fixed stellar mass. Taken together, the turnover in the MZR is the result of the conflict between the trend for more massive galaxies to have higher SFRs and the trend for metallicity to decrease with SFR at fixed mass. The turnover in the T04 MZR (and other strong line calibration MZRs) occurs at a higher stellar mass than the the direct method MZR because the strong line metallicity calibrations produce a weaker
SFR–metallicity anticorrelation. This means that the progression to higher SFRs with increasing stellar mass has less of an effect on the MZR.

Interestingly, the SFR$^{0.0}_{-0.5}$ stacks (light green circles/line) are nearly identical to the T04 MZR in slope, shape, turnover, and normalization. While the exact cause of this agreement is unclear, it is possible that the photoionization models that underlie the T04 metallicities assume physical parameters that are most appropriate for galaxies with this range of SFR. We discuss potential systematic effects of strong line calibrations in Section 2.7.3.

The stacks with very high SFRs (SFR$^{1.5}_{1.0}$ and SFR$^{2.0}_{1.5}$; blue and dark blue circles/lines, respectively) have significantly lower metallicities than the stack of all galaxies at fixed mass in the MZR. The high SFRs and low metallicities of these galaxies suggests that they are probably undergoing major mergers, as found by Peeples et al. (2009) for similar outliers. Major mergers drive in considerable amounts of low metallicity gas from large radii, which dilutes the metallicity of the galaxy and triggers vigorous star formation (e.g., Kewley et al. 2006, 2010; Torrey et al. 2012). These stacks also have a larger scatter than lower SFR stacks, which is likely driven by the small numbers of galaxies per stack coupled with the large intrinsic dispersion in the individual galaxy metallicities.
2.5.3. The Fundamental Metallicity Relation

The orientation of the $M_\star - Z - SFR$ relation captures the importance of SFR as a second parameter to the MZR (Lara-López et al. 2010; Mannucci et al. 2010). Mannucci et al. (2010) established the convention that the FMR is the projection of least scatter found by choosing a free parameter $\alpha$ that minimizes the scatter in the metallicity vs. $\mu_\alpha \equiv \log(M_\star) - \alpha \log(SFR)$ plane (Equation 2.1). Mannucci et al. (2010) found a value of $\alpha = 0.32$ for a sample of SDSS galaxies with metallicities determined with the semi-empirical calibration of Maiolino et al. (2008). As metallicity estimates are well known to vary substantially between different methods, the parameter $\alpha$ may also be different due to potentially different correlations between the inferred metallicity and the SFR. For example, Yates et al. (2012) used the T04 metallicities, rather than those employed by Mannucci et al. (2010), and found a lower value of $\alpha = 0.19$.

Figure 2.12 shows the fundamental metallicity relation for the $M_\star - SFR$ stacks (circles color-coded by SFR). The scatter in metallicity at fixed $\mu_\alpha$ is minimized for $\alpha = 0.66$, which is significantly larger than the $\alpha$ values found by Mannucci et al. (2010) and Yates et al. (2012) for metallicities estimated with strong line calibrations. The scatter for the stacks differs from the scatter for individual galaxies (like the Mannucci et al. 2010 and Yates et al. 2012 studies) because the number of galaxies per stack varies. For a direct comparison, we computed the value of $\alpha$ for the metallicities derived from various empirical, semi-empirical, and theoretical methods.
strong line calibrations for the stacks with log(SFR) > −1.0 (the same SFR range as the stacks with direct method metallicities) and find low α values (α = 0.12–0.34) that are consistent with the Mannucci et al. (2010) and Yates et al. (2012) α values (see Table 5 of Andrews & Martini 2013). The significant difference in α between the direct method and the strong line methods indicates that the calibrations of all of the strong line methods have some dependence on physical properties that correlate with SFR.

The scatter in the direct method FMR (σ = 0.13 dex; Figure 2.12) is almost a factor of two smaller than the scatter for the $M_\ast$–SFR stacks with direct method metallicities in the mass–metallicity plane (σ = 0.22 dex; Figure 2.11). This decrease is due to two features of the $M_\ast$–SFR stacks at fixed SFR shown as the solid colored lines in Figure 2.11: (1) they are substantially offset from one another; (2) they have similar slopes with minimal overlap. The former reflects a strong SFR-dependence on the MZR; the latter corresponds to a monotonic SFR–metallicity relation at fixed stellar mass.

Figure 2.13 shows the $M_\ast$–SFR stacks (circles color-coded by SFR) in the mass–metallicity plane with metallicities determined with two representative strong line calibrations: the Kobulnicky & Kewley (2004) theoretical $R_{23}$ calibration (panel a) and the Pettini & Pagel (2004) empirical N2 calibration (panel b). Only stacks with log($M_\ast$) ≥ 8.0 were included in Figure 2.13 because some stacks at lower stellar masses had unphysically high strong line metallicities; to facilitate comparison with
Figure 2.11 only stacks with log(SFR) > −1.0 are shown in Figure 2.13. The stacks in panel (a) show the metallicity from the upper branch of $R_{23}$, which were selected to have log([N II] $\lambda$6583/Hα) > −1.1 (Kewley & Ellison 2008). Panel (b) shows stacks with $-2.5 < \text{log}([\text{N II]} \lambda 6583/\text{H}\alpha) < -0.3$, the calibrated range for the Pettini & Pagel (2004) N2 calibration according to Kewley & Ellison (2008). For reference, the thick black line shows the direct method MZR. The median, 68% contour, and 95% contour of the T04 MZR are indicated by the solid, dashed, and dotted gray lines, respectively.

The scatter in metallicity about the best fit relation decreases only marginally from the MZR to the FMR when strong line calibrations are used to estimate metallicity. For the KK04 and PP04 N2 metallicities, the scatter is reduced by $\sigma = 0.10 \rightarrow 0.09$ dex and $\sigma = 0.10 \rightarrow 0.07$ dex, respectively. Figure 2.13 also shows that the constant SFR tracks for the strong line calibrations in the mass–metallicity plane are both more closely packed and overlap more than those of the direct method. Figure 2.13 only shows the $M_\ast$–SFR stacks with metallicities from two strong line calibrations, one theoretical and one empirical, but the minor reduction in scatter, small spread, and considerable overlap are generic features of strong line metallicities (the normalization is not).

A qualitative measure of the spread is the difference between the metallicity of the SFR$_{0.0}^{0.0}$ (light green) and the SFR$_{0.5}^{1.0}$ (light blue) stacks at a given stellar mass. There are 17 stellar mass bins with direct method metallicities for stacks with these
SFRs. The median metallicity difference for these pairs of stacks was 0.38 dex for the direct method, 0.15 dex for the Kobulnicky & Kewley (2004) calibration, and 0.13 dex for the Pettini & Pagel (2004) calibration. The factor of ~2–3 difference between the direct method and strong line metallicities translates into an analogous difference in the SFR-dependence of the MZR.

Another feature of the strong line MZRs at fixed SFR is that different SFR tracks turn over at different stellar masses. Low SFR tracks turn over at lower stellar masses than high SFR tracks, so the sign of the dependence of the MZR on SFR changes with stellar mass. At low stellar masses, higher SFR stacks have lower metallicities; at high stellar masses, the opposite is true—higher SFR stacks have higher metallicities. Yates et al. (2012) found a similar, but more dramatic, result for their sample of galaxies that used T04 metallicities. The origin of the weak SFR-dependence and non-monotonic relation for the strong line calibrations is not obvious, but we discuss several potentially relevant effects in Section 2.7.3.

2.6. N/O Abundance

Nitrogen provides interesting constraints on chemical evolution because it is both a primary and secondary nucleosynthetic product. The yields of primary elements are independent of the initial metal content of a star but the yields of secondary elements are not. In a low metallicity star, the majority of the seed carbon and oxygen nuclei that will form nitrogen during the CNO cycle are created during
helium burning in the star, so the nitrogen yield of such a star will scale roughly
with the carbon and oxygen yields. In this case, carbon, nitrogen, and oxygen all
behave like primary elements. After the ISM becomes sufficiently enriched, the
nitrogen yield of a star principally depends on the amount of carbon and oxygen
incorporated in the star at birth. The carbon and oxygen still behave like primary
elements, but nitrogen is a secondary nucleosynthetic product. Observational studies
(Vila Costas & Edmunds 1993; van Zee & Haynes 2006; Berg et al. 2012) have found
clear evidence for primary and secondary nitrogen at low and high metallicity. Vila
Costas & Edmunds (1993) created a simple, closed box chemical evolution model
that quantified the regimes where nitrogen is expected to behave like a primary and
secondary element. However, modeling nitrogen enrichment is difficult because of
the large uncertainties in stellar yields and the delay time for nitrogen enrichment
relative to oxygen. Galactic winds also complicate nitrogen enrichment because they
may preferentially eject oxygen relative to nitrogen (van Zee & Haynes 2006). This
is because oxygen is formed quickly in massive stars and is available to be ejected
from galaxies by winds associated with intense bursts of star formation. By contrast,
the <100 Myr delay before the release of nitrogen from intermediate mass AGB
stars might be sufficient to protect it from ejections by galactic winds.

In principle, the N/O abundance as a function of oxygen abundance can be
used to disentangle the effects of nucleosynthesis, galactic inflows and outflows, and
different star formation histories on the relative enrichment of nitrogen. The total
N/O ratio is a difficult quantity to measure because [N III] lines are not readily observable, so N\(^+\)/O\(^+\) is used frequently as a proxy for N/O. This assumption is supported by the photoionization models of Garnett (1990), which showed that the ionization correction factor from N\(^+\)/O\(^+\) to N/O should be $\sim 1$ to within 20%.

Because the ionization factor should be close to unity, most papers in the literature (e.g., Vila Costas & Edmunds 1993) that show N/O have assumed N/O = N\(^+\)/O\(^+\).

For transparency, we plot N\(^+\)/O\(^+\) as a function of direct method oxygen abundance in Figure 2.14a for the $M_*$ and $M_*$–SFR stacks. We measured the ionic abundances of N\(^+\) and O\(^+\) with the direct method under the assumption that $T_e$[O II] represents $T_2$ (see Section 2.3.2).

At low metallicity ($12 + \log(O/H) < 8.5$), we find that the $M_*$ stacks have an approximately constant value of N\(^+\)/O\(^+\), which is expected for primary nitrogen. These stacks have a median of $\log(N^+/O^+) = -1.43$ (indicated by the horizontal line[3]), which is consistent with other studies of H II regions and dwarf galaxies (Vila Costas & Edmunds 1993). At $12 + \log(O/H) = 8.5$, there is a sharp transition where N\(^+\)/O\(^+\) increases steeply with oxygen abundance (slope = 1.73), which shows that nitrogen is acting like a secondary element. Previous observations (e.g., Vila Costas & Edmunds 1993) have found a smoother transition between primary and

---

[3] We do not show a fit to these points because of the strong a priori expectation of a constant N\(^+\)/O\(^+\) at low metallicity (and low mass); however, a linear fit would have a slope of $-0.21$. The analogous slope for the low mass N\(^+\)/O\(^+\)–$M_*$ relation is $-0.08$.  

52
secondary nitrogen and a shallower slope in the secondary nitrogen regime, albeit with large dispersion that could be obscuring these features. The fit parameters of the $N^+/O^+–O/H$ relation for the $M_*$ stacks is presented in Table 6 of Andrews & Martini (2013).

The $M_*$ stacks form a tight sequence with a dispersion of only $\sigma = 0.08$ dex, compared to a more typical dispersion of $\sigma \sim 0.3$ dex for individual objects (e.g., Henry et al. 2000). A plausible explanation for the additional scatter in the N/O–O/H relation for individual galaxies is the time-dependence of N/O caused by the difference in enrichment timescales of oxygen and nitrogen following a burst of star formation. The $M_*$–SFR stacks show a larger dispersion than the $M_*$ stacks, potentially because these stacks contain fewer galaxies. The low and moderate SFR stacks ($\text{SFR}^{-0.5}_{-1.0}$, $\text{SFR}^{0.0}_{-0.5}$, and $\text{SFR}^{0.5}_{0.0}$) follow the general trend of the $M_*$ stacks; however, the high SFR stacks ($\text{SFR}^{1.0}_{0.5}$, $\text{SFR}^{1.5}_{1.0}$, and $\text{SFR}^{2.0}_{1.5}$) have higher $N^+/O^+$ at a given oxygen abundance, which may be because these galaxies have experienced a large inflow of gas that would lower O/H at fixed N/O (i.e., move galaxies to the left in Figure 2.14a). Another consequence of a vigorous burst of star formation is the production of Wolf-Rayet stars that can enrich the gas in nitrogen for a brief period before the oxygen enrichment from the subsequent SNe II (Berg et al. 2011). We see evidence for Wolf-Rayet features, such as He II $\lambda 4686$, in some of our stacks, especially at low mass.
Some of the features in the \( \text{N}^+ / \text{O}^+ - \text{O}/\text{H} \) relation are clarified by the associated relation between \( \text{N}^+ / \text{O}^+ \) and stellar mass, which is shown in Figure 2.14b for the \( M_* \) and \( M_* - \text{SFR} \) stacks (see Table 6 of Andrews & Martini 2013 for the fit parameters of the \( \text{N}^+ / \text{O}^+ - M_* \) relation for the \( M_* \) stacks). Similar to Figure 2.14a, there is a primary nitrogen plateau in \( \text{N}^+ / \text{O}^+ \) at low stellar mass (\( \log[M_*] < 8.9 \)) and a steady increase in \( \text{N}^+ / \text{O}^+ \) due to secondary nitrogen enrichment above \( \log(M_*) = 8.9 \) (slope = 0.30). However, in the secondary nitrogen regime, the \( \text{N}^+ / \text{O}^+ - M_* \) relation has a much lower dispersion (\( \sigma = 0.01 \) dex) than the \( \text{N}^+ / \text{O}^+ - \text{O}/\text{H} \) relation (\( \sigma = 0.08 \) dex). Some of the decreased dispersion is due to the larger dynamic range of stellar mass relative to oxygen abundance, but the tightness of the \( \text{N}^+ / \text{O}^+ - M_* \) relation suggests that the enrichment of nitrogen relative to oxygen is well-behaved on average. The essentially zero intrinsic dispersion in the \( \text{N}^+ / \text{O}^+ - M_* \) relation can be used to quantify the effect of gas inflow and galactic winds on enrichment if all of the scatter in \( \text{N}^+ / \text{O}^+ \) at a given \( \text{O}/\text{H} \) is due to gas flows into and out of galaxies. As in the \( \text{N}^+ / \text{O}^+ - \text{O}/\text{H} \) relation, the low and moderate SFR stacks (\( \text{SFR}_{-0.5}^{0.5}, \text{SFR}_{-0.5}^{0.0}, \) and \( \text{SFR}_{0.0}^{0.5} \)) roughly coincide with the \( M_* \) stacks. The high SFR stacks (\( \text{SFR}_{0.5}^{1.0}, \text{SFR}_{1.0}^{1.5}, \) and \( \text{SFR}_{1.5}^{2.0} \)) still tend to be more nitrogen enriched than the \( M_* \) stacks at a fixed \( M_* \), but the discrepancy has decreased. The \( \text{N}^+ / \text{O}^+ - M_* \) diagram is less sensitive to dilution (traced by SFR) because the high SFR galaxies with low \( \text{O}/\text{H} \) are less significant outliers when shown as a function of stellar mass.
The N/O–M$_*$ relation has been previously investigated by Pérez-Montero & Contini (2009) and Pérez-Montero et al. (2013), who used strong line methods to estimate N/O. They found that N/O increased steadily with stellar mass and did not show a plateau at low stellar mass associated with primary nitrogen enrichment, in contrast to the direct method N/O–M$_*$ relation. However, Pérez-Montero et al. (2013) showed that the strong line N/O–M$_*$ relation is nearly independent of SFR, which is roughly consistent with our finding that the N/O–M$_*$ relation has only a mild dependence on SFR, particularly at log($M_*$) $\gtrsim$ 9.0.

2.7. Discussion

2.7.1. Comparison to a Previous Analysis That Used Auroral Lines from Stacked Spectra

Liang et al. (2007) stacked SDSS spectra and applied the direct method to estimate the MZR, although their study differs from ours in a number of important respects. First, their study is based on DR4 spectroscopy of 23,608 galaxies, which is approximately an order of magnitude fewer than our sample. Second, they implemented a minimum [O II] $\lambda$3727 EW criterion to select the input galaxies to their stacks in order to increase the SNR of their stacked spectra. Finally, they only measured $T_e$[O II] from the [O II] $\lambda\lambda$7320, 7330 lines and then inferred $T_e$[O III] (and the O$^{++}$ ionic abundance) from the $T_2$–$T_3$ relation provided by Izotov et al. (2006).
These differences are likely responsible for the offset between our MZR and the Liang et al. (2007) MZR, the absence of a turnover in their MZR, and their greater scatter as shown in Figure 2.15.

The \([\text{O} \, \text{II}]\) selection criterion can readily explain part of the offset between our MZRs. Liang et al. (2007) only selected galaxies with above average \([\text{O} \, \text{II}]\) \(\lambda 3727\) EW (at fixed mass) for galaxies with \(\log(M_\star) < 10\) and required a more stringent \(\text{EW}([\text{O} \, \text{II}]) > 30 \, \text{Å}\) for galaxies with \(\log(M_\star) > 10\). As a result of this selection, their stacks have systematically higher SFRs by approximately 0.15 to 0.2 dex. This in turn biases the stacks to lower metallicities because of the \(M_\star-Z-SFR\) relation (Mannucci et al. 2010; Lara-López et al. 2010). The magnitude of this effect (\(\sim 0.05-0.08\) dex) accounts for part of the difference between the MZRs. Another effect of this selection is that the increase in average SFR increases the turnover mass and makes it less distinct (see Figure 2.11).

The turnover mass is also not apparent in their MZR due to the greater scatter, which is largely due to their order of magnitude smaller sample. The scatter around the linear fit from \(\log(M_\star) = 8.0-10.5\) for their data is \(\sigma = 0.12\) dex. The scatter around an asymptotic logarithmic fit (Equation 2.5) is reduced only to \(\sigma = 0.11\) dex. An asymptotic logarithmic fit has an additional degree of freedom relative to a linear fit, so the marginal improvement in \(\sigma\) suggests that the Liang et al. (2007) MZR can be sufficiently characterized by a linear fit. Over the same mass range, the scatter around the asymptotic logarithmic fit of our data (thick black line) is only
\( \sigma = 0.03 \) dex, or a factor of four smaller. The smaller scatter in our MZR enables a clear identification of the turnover.

The method employed by Liang et al. (2007) to estimate the oxygen abundance is also distinct from ours and may explain the rest of the discrepancy in the normalization difference between our studies. The Liang et al. (2007) study relies solely on the [O II] \( \lambda\lambda 7320, 7330 \) auroral lines to measure \( T_e[\text{O II}] \), which is used to infer \( T_e[\text{O III}] \) and the \( \text{O}^{++} \) abundance by applying the \( T_2-T_3 \) relation and \( T_e[\text{O III}]-(\text{O}^{++}/\text{H}^+) \) formula from Izotov et al. (2006). They did not detect [O III] \( \lambda 4363 \) in their stacks, which they only binned in stellar mass, because they had fewer galaxies per bin. The stellar continuum subtraction may also have affected the detection of [O III] \( \lambda 4363 \) because of its proximity to the H\( \gamma \) stellar absorption feature, whereas the stellar continuum is comparatively featureless in the vicinity of the [O II] \( \lambda\lambda 7320, 7330 \) lines. Liang et al. (2007) used the Bruzual & Charlot (2003) spectral templates, rather than the empirical and higher resolution MILES templates that we have adopted (see Section 2.2.3), and this difference may also have played an important role. As a consequence of their lack of a detection of [O III] \( \lambda 4363 \), their oxygen abundance estimate depends on the quality of the assumption that the galaxies obey the \( T_2-T_3 \) relation of Izotov et al. (2006). Our empirical measurements of \( T_2 \) and \( T_3 \) indicate that this assumption underestimates \( T_3 \) and overestimates \( \text{O}^{++}/\text{H}^+ \), which may partly explain why our MZRs are in better agreement at high mass where \( \text{O}^+ \) is the dominant ionization state of oxygen.
2.7.2. Temperature and Metallicity Discrepancies

Temperatures and metallicities of H II regions measured with the direct method do not always agree with those measured with other techniques. For example, temperatures measured with the direct method tend to be systematically higher than those measured from the Balmer continuum (Peimbert 1967). Also, the metallicities determined from optical recombination lines (e.g., C II λ4267 and O II λ4649) and far-IR fine-structure lines (e.g., [O III] 52, 88 μm) tend to be 0.2–0.3 dex higher than those from collisionally excited lines (García-Rojas & Esteban 2007; Bresolin 2008; Esteban et al. 2009). The exact cause of these temperature and abundance discrepancies is currently not understood.

Peimbert (1967) proposed that temperature fluctuations and gradients in H II regions cause direct method temperatures to be systematically overestimated, while direct method metallicities are underestimated. To account for temperature variations across a nebula, he introduced the concept of $t^2$, the root mean square deviation of the temperature from the mean. Estimating $t^2$ has proven to be difficult, so most direct method metallicity studies assume $t^2 = 0$. However, optical recombination lines and far-IR fine-structure lines (Garnett et al. 2004a) are less sensitive to temperature than collisionally excited lines, so they could be used to estimate $t^2$ if the discrepancy between the metallicity determined from collisionally excited lines and optical recombination lines or far-IR fine-structure lines is assumed to be caused by temperature fluctuations. The few studies that have measured
optical recombination lines (e.g., García-Rojas & Esteban 2007; Esteban et al. 2009) find that values of $t^2 = 0.03$–0.07 are necessary to increase the direct method metallicities by 0.2–0.3 dex to match the optical recombination line metallicities.

Recently, Nicholls et al. (2012) suggested that the electron energy distribution could be the cause of the temperature and metallicity discrepancies. Specifically, they questioned the widespread assumption that the electrons are in thermal equilibrium and can be described by a Maxwell-Boltzmann distribution. Instead, they suggested that there might be an excess of high energy electrons and proposed that a $\kappa$-distribution is a more appropriate description of the electron energy distribution. The $\kappa$-distribution is based on direct measurements of solar system plasmas. Assuming a $\kappa$-distribution for an H II region lowers the derived temperature, increases the inferred metallicity, and could potentially resolve the discrepancy between the temperatures and metallicities found with optical recombination lines and collisionally excited lines.

Models of H II regions by Stasińska (2005) indicate that metallicities based on the direct method could suffer from systematic biases in metal-rich H II regions. She finds that measuring metallicity from $T_e[O \text{ III}]$ and $T_e[N \text{ II}]$ tends to dramatically underestimate the true metallicity for $12 + \log(O/H) > 8.6$ (see her Figure 1). The situation does not improve if metallicities are computed with only $T_e[N \text{ II}]$ because the derived metallicity can wildly overestimate or underestimate the true metallicity depending on the physical conditions and geometry of the H II region. However,
there are two key differences between the models of Stasińska (2005) and the measurements made in this study that could minimize the bias. First, we measured the temperature of the low ionization region from $T_e[O\text{II}]$, not $T_e[N\text{II}]$. Second, we analyzed spectra of galaxy stacks and not individual H II regions, which may average out the large predicted errors. Our stacks generally increase smoothly in metallicity as a function of $M_*$, which does not rule out systematic error but minimizes the impact of the individual, catastrophic errors highlighted in Stasińska (2005).

2.7.3. Strong Line Calibrations and the SFR-dependence of the FMR

The MZRs and FMRs based on strong line calibrations (Figure 2.13) have a much weaker dependence on SFR than the direct method MZR and FMR (Figures 2.10–2.12). Relative to the direct method MZR and FMR, the strong line MZR and FMR have (1) a smaller spread in the mass–metallicity plane (compare Figures 2.11 and 2.13), (2) a smaller reduction in scatter from the MZR to the FMR (see Section 2.5.3), and (3) a smaller value of $\alpha$ (see Table 5 of Andrews & Martini 2013). This trend is a generic feature of strong line calibrations that holds for both empirical and theoretical calibrations and for all strong line indicators ($R_{23}$, N2, N2O2, and O3N2) that we used. Since a strong line calibration is only applicable to the physical conditions spanned by the calibration sample or model, it is important to understand the physical properties of the calibration sample for empirical calibrations and the assumptions behind the H II region models that underlie theoretical calibrations.
Figure 2.16 compares excitation parameter ($P$) and $R_{23}$ (panel a) and [O II] and [O III] fluxes relative to H$\beta$ (panel b) for galaxies, stacks of galaxies, and H II regions. The gray contours (50%, 75%, and 95%) and points indicate SDSS star-forming galaxies, whose line flux measurements come from the MPA-JHU catalog (Tremonti et al. 2004), after we corrected their measured values for intrinsic reddening. The stacks are shown by the open and colored circles. Extragalactic H II regions with direct method metallicities are represented by the light blue contours and crosses.

The dereddened line fluxes of the H II regions come from the literature compilation by Pilyugin et al. (2012).4

Figure 2.16a shows the excitation parameter $P$ as a function of $R_{23}$. Excitation increases upwards, but $R_{23}$ is double-valued with metallicity, so metallicity increases to the left for objects on the upper branch (the majority of the galaxies and stacks) and increases to the right for objects on the lower branch (most of the compiled H\,\II regions). Figure 2.16b displays another projection of the same data in the space defined by the dereddened [O\,\II] $\lambda 3727$ and [O\,\III] $\lambda 5007$ line fluxes relative to H$\beta$. The dotted lines show constant $R_{23}$ values, and the dashed lines mark constant [O\,\II] $\lambda 3727$/[O\,\III] $\lambda 5007$ values.

In Figure 2.16, the compiled H\,\II regions predominantly overlap with the high excitation and high $R_{23}$ tail of the galaxy distribution in Figure 2.16a and the analogous high [O\,\III] $\lambda 5007$ tail of the galaxy distribution in Figure 2.16b, which corresponds to low metallicity galaxies. The compiled H\,\II regions have direct method metallicities and therefore at least one detectable auroral line, usually [O\,\III] $\lambda 4363$. Because the strength of the auroral lines, especially [O\,\III] $\lambda 4363$, is

a strong function of metallicity and excitation parameter, these H II regions were effectively selected to have low metallicities and high excitation parameters. Thus, they are not representative of the typical conditions found in the H II regions of the galaxy sample. Empirical calibrations, which are based on samples of H II regions with direct method metallicities, are not well constrained in the high metallicity, low excitation regime where most galaxies and their constituent H II regions lie. For example, Moustakas et al. (2010) recommended only using the empirical Pilyugin & Thuan (2005) $R_{23}$ calibration for objects with $P > 0.4$. When empirical calibrations are applied to large galaxy samples, galaxy metallicities are systematically underestimated, particularly at low excitation and high metallicity (Moustakas et al. 2010). Similarly, MZRs based on empirical calibrations may have an artificially weak dependence on SFR.

The typical excitation conditions and $R_{23}$ values of the stacks are much better matched to the overall galaxy distribution than the compiled H II regions with direct method metallicities. The stacks probe to both lower excitation ($P \approx 0.2$) and higher metallicity ($R_{23} \approx 0.4$) than the bulk of the compiled H II regions. The stacks do not continue to low $R_{23}$ values ($<0.3$), a region of parameter space populated by the most massive and metal-rich galaxies in our sample. The [O II] $\lambda 3727$ and [O III] $\lambda 5007$ line fluxes of these galaxies vary significantly, even at the same stellar mass and SFR. While the stacks do not reach the lowest $R_{23}$ values of the galaxies, they still trace the average $R_{23}$ values of the galaxies in each stack.
Theoretical calibrations are based on stellar population synthesis models, like STARBURST99 (Leitherer et al. 1999), and photoionization models, such as MAPPINGS (Sutherland & Dopita 1993; Groves et al. 2004a,b) and CLOUDY (Ferland et al. 1998). The stellar population synthesis model generates an ionizing radiation field that is then processed through the gas by the photoionization model. The parameters in the stellar population synthesis model include stellar metallicity, age of the ionizing source, initial mass function, and star formation history. In the photoionization model, the electron density and the ionization parameter are adjustable parameters. Because the model grids can span a wide range of parameter space, particularly in metallicity and excitation parameter, theoretical calibrations have an advantage over empirical calibrations at high metallicity and low excitation, where empirical calibrations are not strictly applicable.

However, metallicities derived with theoretical calibrations can be significantly higher (up to 0.7 dex; see Kewley & Ellison 2008) than direct method metallicities. The most likely cause of this offset is the breakdown of one or more of the assumptions about the physics of H II regions in the stellar population synthesis or photoionization models. In the stellar population synthesis models, the ionizing source is usually treated as a zero age main sequence starburst, which is not applicable for older star clusters (Berg et al. 2011), and the line fluxes can change appreciably as a cluster (and the associated H II region) ages. As elucidated by Kewley & Ellison (2008), there are three main issues with the photoionization
models. First, they treat the nebular geometry as either spherical or plane-parallel, which may not be appropriate for the true geometries of the H II regions. Second, the fraction of metals depleted onto dust grains is poorly constrained by observations (see Draine 2003; Jenkins 2009) but is a required parameter of the photoionization models. Third, they assume that the density distribution of the gas and dust as smooth, when it is clumpy. While all these assumptions might break down to some degree, it is unknown which assumption or assumptions causes metallicities based on theoretical strong line calibrations to be offset from the direct method metallicities, but it is conceivable that the weak SFR dependence of theoretical strong line calibration MZRs is also due to these assumptions.

One of the most intriguing findings of the Mannucci et al. (2010) and Lara-López et al. (2010) studies is that high redshift observations are consistent with no redshift evolution of the strong line FMR up to \(z = 2.5\) and \(z = 3.5\), respectively. Given the large discrepancies between the local strong line and direct method FMRs, a fair comparison between the local direct method FMR and a high redshift strong line FMR is not possible. An interesting test would be to check if high redshift direct method metallicity measurements are consistent with the local direct method FMR. A few studies (Hovos et al. 2005; Kakazu et al. 2007; Yuan & Kewley 2009; Erb et al. 2010; Brammer et al. 2012) have reported direct method metallicities at higher redshifts (\(z \sim 0.7–2.3\)), but none simultaneously provide the stellar masses and SFRs of the galaxies. Since the FMR and its evolution provide important constraints on
theoretical galaxy evolution models and form the basis of empirical galaxy evolution models (Zahid et al. 2012b; Peeples & Somerville 2013), future studies that measure all three of these parameters would be valuable.

2.7.4. Physical Processes Governing the MZR and $M_*-Z-SFR$ Relation

Understanding the baryon cycling of galaxies relies heavily on the adopted relations between stellar mass, metallicity, and SFR. Traditionally, the MZR and $M_*-Z-SFR$ relation have been measured with strong line methods. In this study, we have used the more reliable direct method to measure the MZR and $M_*-Z-SFR$ relation. The direct method MZR (Figure 2.10) spans three orders of magnitude in stellar mass from $\log(M_*) = 7.4-10.5$ and thus simultaneously extends the MZR to lower masses by an order of magnitude compared to strong line MZRs (e.g., T04) and resolves the high mass turnover. The features of the direct method MZR that most strongly influence the physical interpretations are its low mass slope ($O/H \propto M_*^{1/2}$), its turnover mass ($\log[M_*] = 8.9$), and its normalization ($12 + \log(O/H)_{asm} = 8.8$). The SFR-dependence of the MZR (see Figures 2.11 and 2.12) also serves as an important observational constraint for galaxy evolution models. We find that the MZR depends strongly on SFR ($\alpha = 0.66$; Figure 2.12) at all stellar masses.

The MZR and $M_*-Z-SFR$ relation are shaped by gas inflows, gas outflows, and star formation. The interplay between these three processes is complex, so
hydrodynamic galaxy simulations (e.g., Brooks et al. 2007; Finlator & Davé 2008; Davé et al. 2011b; Davé et al. 2011a) and analytic models (e.g., Peeples & Shankar 2011; Davé et al. 2012) have been used to establish a framework to interpret the observations in a physical context. Below we briefly discuss the physical implications of our results within the formalisms of Peeples & Shankar (2011) and Finlator & Davé (2008).

Peeples & Shankar (2011) developed an analytic model for understanding the importance of outflows in governing the MZR based on the assumption that galaxies follow zero scatter relations between stellar mass, gas fraction, metallicity, outflow efficiency, and host halo properties. In their formalism, the primary variable controlling the MZR is the metallicity-weighted mass-loading parameter,

$$\zeta_{\text{wind}} \equiv \left( \frac{Z_{\text{wind}}}{Z_{\text{ISM}}} \right) \left( \frac{\dot{M}_{\text{wind}}}{\dot{M}_*} \right),$$

(2.6)

where $Z_{\text{wind}}$ and $Z_{\text{ISM}}$ are the wind and ISM metallicities, respectively, and $\dot{M}_{\text{wind}}/\dot{M}_*$ is the unweighted mass-loading parameter. $\zeta_{\text{wind}}$ can be expressed in terms of the MZR and the stellar mass–gas fraction relation by rearranging their Equation (20):

$$\zeta_{\text{wind}} = y/Z_{\text{ISM}} - 1 - \alpha F_{\text{gas}},$$

(2.7)

where $y$ is the nucleosynthetic yield, $\alpha$ is a parameter of order unity (see their Equation 11), and $F_{\text{gas}} \equiv M_{\text{gas}}/M_*$ is the gas fraction.
If we adopt the Peeples & Shankar (2011) formalism and their fiducial yield and stellar mass–gas fraction relation, then we can solve for the $M_\star$–$\zeta_{\text{wind}}$ relation implied by the direct method MZR. This direct method $M_\star$–$\zeta_{\text{wind}}$ relation starts at high $\zeta_{\text{wind}}$ ($\zeta_{\text{wind}} \sim 15$) for low mass galaxies ($\log[M_\star] = 7.5$). Then, $\zeta_{\text{wind}}$ decreases with increasing stellar mass, eventually flattening and approaching a constant $\zeta_{\text{wind}}$ ($\zeta_{\text{wind}} \sim 2$) above the turnover mass ($\log[M_\star] = 8.9$). Since the D02 MZR has a similar shape and normalization to the direct method MZR from $\log(M_\star) = 8.5$–10.5, the direct method $M_\star$–$\zeta_{\text{wind}}$ relation resembles the D02 $M_\star$–$\zeta_{\text{wind}}$ relation shown in Figure 6 of Peeples & Shankar (2011). Also, the direct method MZR implies a similar behavior for $Z_{\text{wind}}$ and $Z_{\text{ISM}}$ as a function of stellar mass as the D02 MZR (see their Figure 9). The ratio of $Z_{\text{wind}}/Z_{\text{ISM}}$ inversely correlates with how efficiently winds entrain ambient ISM. If we adopt the simple relation between metallicity and the unweighted mass-loading parameter from Finlator & Davé (2008), $Z_{\text{ISM}} \approx y/(1 + \dot{M}_{\text{wind}}/\dot{M}_\star)$, then the direct method MZR implies an efficiency of mass ejection that scales as $\dot{M}_{\text{wind}}/\dot{M}_\star \propto M_\star^{-1/2}$ for $\log(M_\star) \lesssim 9.0$. The higher $\zeta_{\text{wind}}$ for low mass galaxies relative to high mass galaxies could be due to more enriched winds (larger $Z_{\text{wind}}/Z_{\text{ISM}}$) or more efficient mass ejection by winds (larger $\dot{M}_{\text{wind}}/\dot{M}_\star$) or both. Peeples & Shankar (2011) found that the $M_\star$–$\zeta_{\text{wind}}$ relation follows the general shape of the direct method $M_\star$–$\zeta_{\text{wind}}$ relation regardless of the input MZR (see their Figure 6). However, the direct method MZR requires more efficient metal ejection by winds than theoretical strong line calibration MZRs (T04; Z94; KK04; M91) at
all stellar masses because of the lower normalization of the direct method MZR. We note that the yield is poorly constrained, and a higher adopted yield requires more efficient outflows to produce the observed MZR.

In contrast to the Peeples & Shankar (2011) framework that assumed a zero scatter MZR (and therefore does not account for variations in the SFR or gas fraction at a fixed stellar mass), the Finlator & Davé (2008) model, based on cosmological hydrodynamic simulations, treats the MZR as an equilibrium condition. In their model, galaxies are perturbed off the MZR by stochastic inflows but the star formation triggered by the inflow of gas and the subsequent metal production returns them to the mean MZR. The rate at which galaxies re-equilibrate following an episode of gas inflow sets the scatter in the MZR, which is indirectly traced by the SFR-dependence of the $M_*-Z$–SFR relation.

The observed SFR-dependence of the $M_*-Z$–SFR relation differs according to the strong line metallicity calibration used to construct the $M_*-Z$–SFR relation, as found by Yates et al. (2012). Specifically, they used metallicities estimated with the Mannucci et al. (2010) method and T04 method. At low stellar masses, metallicity decreases with increasing SFR for both $M_*-Z$–SFR relations. But at high stellar masses ($\log[M_*] > 10.5$), the SFR-dependence of the T04 $M_*-Z$–SFR relation reverses, so that metallicity increases with increasing SFR; however, the Mannucci et al. (2010) $M_*-Z$–SFR relation collapses to a single sequence that is independent of SFR. Yates et al. (2012) suggested that the SFR-dependence of the Mannucci
et al. (2010) $M_\star$–$Z$–SFR relation at high stellar mass is obscured by the N2 indicator (which was averaged with the metallicity estimated from $R_{23}$) used in the Mannucci et al. (2010) metallicity calibration, which saturates at high metallicity.

Unlike the Mannucci et al. (2010) and T04 $M_\star$–$Z$–SFR relations, the SFR-dependence of the direct method $M_\star$–$Z$–SFR relation does not change dramatically with stellar mass. There is little overlap between the constant SFR tracks in the direct method $M_\star$–$Z$–SFR relation (Figure 2.11). Furthermore, the SFR-dependence is strong ($\alpha = 0.66$; see Section 2.5.3), so the scatter in the direct method MZR for individual galaxies (if it could be measured) would be larger than the scatter in the Mannucci et al. (2010) and T04 MZRs. Within the context of the Finlator & Davé (2008) model, this means that the direct method MZR implies a longer timescale for galaxies to re-equilibrate than the Mannucci et al. (2010) and T04 MZRs. We note that the direct method $M_\star$–$Z$–SFR relation does not probe above log($M_\star$) = 10.5 because the auroral lines are undetected in this regime; however, this mass scale is where the discrepancies between the Mannucci et al. (2010) and T04 metallicities are the largest—potentially due to a break down of strong line calibrations at high metallicities (see Section 2.7.3).

2.8. Summary

We have measured [O III], [O II], [N II], and [S II] electron temperatures, direct method gas-phase oxygen abundances, and direct method gas-phase nitrogen to
oxygen abundance ratios from stacked galaxy spectra. We stacked the spectra of \( \sim 200,000 \) SDSS star-forming galaxies in bins of (1) 0.1 dex in stellar mass and (2) 0.1 dex in stellar mass and 0.5 dex in SFR. The high SNR stacked spectra enabled the detection of the temperature-sensitive auroral lines that are essential for metallicity measurements with the direct method. Auroral lines are weak, especially in massive, metal-rich objects, but we detect [O \text{ III}] \( \lambda 4363 \) up to \( \log(M_\star) = 9.4 \) and [O \text{ II}] \( \lambda \lambda 7320, 7330 \) up to \( \log(M_\star) = 10.5 \), which is generally not feasible for spectra of individual galaxies. We used the auroral line fluxes to derive the [O \text{ III}] and [O \text{ II}] electron temperatures, the O\text{ ++} and O\text{ +} ionic abundances, and the total oxygen abundances of the stacks.

We constructed the direct method mass–metallicity and \( M_\star-Z-SFR \) relations across a wide range of stellar mass (\( \log(M_\star) = 7.4-10.5 \)) and SFR (\( \log([SFR] = -1.0-2.0) \)). The direct method MZR rises steeply (\( O/H \propto M_\star^{1/2} \)) from \( \log(M_\star) = 7.4-8.9 \). The direct method MZR turns over at \( \log(M_\star) = 8.9 \), in contrast to strong line MZRs that typically turn over at higher masses (\( \log(M_\star) \sim 10.5 \)). Above the turnover, the direct method MZR approaches an asymptotic metallicity of \( 12 + \log(O/H) = 8.8 \), which is consistent with empirical strong line calibration MZRs but \( \sim 0.3 \) dex lower than theoretical strong line calibration MZRs like the Tremonti et al. (2004) MZR. Furthermore, we found that the SFR-dependence (as measured by the value of \( \alpha \) that minimizes the scatter at fixed \( \mu_\alpha \equiv \log(M_\star) - \alpha \log(SFR) \) in the fundamental metallicity relation; see Equation 2.1) of the direct method \( M_\star-Z-SFR \)
relation is \(\sim 2-3\) times larger \((\alpha = 0.66)\) than for strong line \(M_{*} - Z - SFR\) relations \((\alpha \sim 0.12-0.34)\). Its SFR-dependence is monotonic as a function of stellar mass, so constant SFR tracks do not overlap, unlike strong line \(M_{*} - Z - SFR\) relations.

We also showed that the direct method N/O relative abundance correlates strongly with oxygen abundance and even more strongly with stellar mass. N/O exhibits a clear transition from primary to secondary nitrogen enrichment as a function of oxygen abundance and stellar mass.

The slope, turnover, normalization, and SFR-dependence of the MZR act as critical constraints on galaxy evolution models and are best measured by methods that do not rely on strong line diagnostics, such as the direct method. Future work should aim to construct a direct method MZR of individual galaxies with high SNR optical spectra that enable the detection of auroral lines in high mass and high metallicity objects. Furthermore, metallicities based on \textit{Herschel Space Observatory} \cite{Pilbratt2010} measurements of the far-IR fine-structure lines \cite{Croxall2012} from the KINGFISH survey \cite{Kennicutt2011} will provide a valuable check on the absolute abundance scale \cite{Garnett2004a}, which is a major outstanding uncertainty for galaxy evolution studies. These types of investigations will improve our understanding of the galaxy formation process, particularly the cycling of baryons between galaxies and the IGM.
2.9. Figures
Fig. 2.1.— Number of galaxies and direct method metallicity as a function of $M_*$ and SFR. The squares represent each $M_*$–SFR stack, the number of galaxies is indicated by the white text, and the color scale corresponds to the metallicity. For reference, the Tremonti et al. (2004) MZR covers $\log(M_*) = 8.5–11.5$, and the Mannucci et al. (2010) FMR spans $\log(M_*) = 9.1–11.35$ and $\log(SFR) = -1.45\rightarrow0.80$.

Fig. 2.2.— Sample spectra from the $\log(M_*) = 8.7–8.8$ ($N_{\text{gal}} = 884$) stack. From left to right, the three columns show the $[\text{O III}] \lambda 4363$, $[\text{N II}] \lambda 5755$, and $[\text{O II}] \lambda\lambda 7320, 7330$ auroral lines. From top to bottom, the four rows correspond to the reduced spectrum of a single galaxy, the spectrum of the stack, the spectrum of the stack after the removal of the stellar continuum (fit from 3700–7360 Å), and the spectrum of the stack after the removal of the stellar continuum (fit to a 200 Å window near the emission line of interest). The continuum rms of each spectrum near the relevant emission line is given in the inset of each panel.
Fig. 2.3.— Electron temperatures derived from the [O II], [N II], and [S II] line ratios plotted as a function of electron temperature derived from the [O III] line ratio for the \( M_\star \) stacks (left column) and \( M_\star \)–SFR stacks (right column; color-coded by SFR). The lines in the top, middle, and bottom rows show the \( T_e\)\([\text{O II}] - T_e\)\([\text{O III}] \), \( T_e\)\([\text{N II}] - T_e\)\([\text{O III}] \), and \( T_e\)\([\text{S II}] - T_e\)\([\text{O III}] \) relations (Equation 2.3), respectively. The outlier in the lower right panel is a single galaxy, so it may not be representative of all galaxies with this stellar mass and star formation rate.
Fig. 2.4.— Electron temperatures derived from the [O II] line ratios as a function of electron temperature derived from the [N II] line ratios for the $M_\star$–SFR stacks (color-coded by SFR). The line indicates $T_{e[O \text{ II}]} = T_{e[N \text{ II}]}$ (as assumed in Equation 2.3).
Fig. 2.5.— The ionic abundance of O\(^+\) (panel a), the ionic abundance of O\(^{++}\) (panel b), and the relative ionic abundance of O\(^{++}\) and O\(^+\) (panel c) as a function of stellar mass for the \(M_\star\) stacks (open circles) and \(M_\star\)-SFR stacks (circles color-coded by SFR). The dashed line in panel (c) indicates equal abundances of O\(^{++}\) and O\(^+\).
Fig. 2.6.— Panel (a): the difference in the direct method metallicity determined from $T_e[\text{O II}]$ only ($T_e[\text{O III}]$ was inferred with the $T_2-T_3$ relation given in Equation 2.3) and the direct method metallicity determined from both $T_e[\text{O II}]$ and $T_e[\text{O III}]$. The dotted line denotes the median difference, and the dashed line marks the upper mass cutoff for which $T_e[\text{O III}]$ can be independently measured in the $M_*$ stacks. Panel (b): the mass-metallicity relation for direct method metallicities determined from $T_e[\text{O II}]$ only (gray circles) and from both $T_e[\text{O II}]$ and $T_e[\text{O III}]$ (open circles). To account for the overestimated metallicity (and underestimated $T_e[\text{O III}]$) caused by assuming the $T_2-T_3$ relation (Equation 2.3), we subtract the median metallicity difference shown in panel (a) from the $T_e[\text{O II}]$-based metallicities above $\log(M_*) = 9.4$ (shown by the dashed line), which results in the open squares. The arrow marks this shift. The sequence of open circles and squares shows the composite direct method metallicities of the $M_*$ stacks that we will adopt for the rest of the paper. We repeated the same procedure for each SFR bin of the $M_*$–SFR stacks. The median metallicity differences are given in Table 4 of Andrews & Martini (2013).
Fig. 2.7.— $T_e[\text{O~III}]$, $T_e[\text{O~II}]$, and direct method metallicity for individual spectra (small gray circles) and stacks in bins of 0.1 dex in stellar mass (large black circles) for the Pilyugin et al. (2010) sample relative to the mean of galaxies within a stellar mass bin of width 0.1 dex $M_\odot$ (shown by the dashed line in each panel). The stacks are consistent with the mean $T_e[\text{O~III}], T_e[\text{O~II}], \text{and metallicity within the measurement uncertainties.}$
Fig. 2.8.— The electron temperatures derived from the [O II], [N II], and [S II] line ratios as a function of electron temperature derived from the [O III] line ratio for the Pilyugin et al. (2010) sample of galaxies with detectable [O III] λ4363 and [O II] λλ7320, 7330 (squares color-coded by SFR; see Section 2.4) and stacks of the same galaxies in bins of 0.1 dex in stellar mass (black circles). The black line shows the $T_2$–$T_3$ relation (Equation 2.3).
Fig. 2.9.— [O II] $\lambda$3727 and [O III] $\lambda$5007 fluxes relative to H$\beta$ of galaxies in one $M_*$–SFR bin ($\log(M_*) = 10.5$–10.6 and $\log([SFR]) = 1.0$–1.5) and the stack of those galaxies. The small black and blue circles represent individual galaxies with fluxes measured with our pipeline and the MPA-JHU pipeline (T04), respectively. The large green circle corresponds to the stack of the same galaxies.
Fig. 2.10.— The direct method mass–metallicity relation for the $M_\star$ stacks (circles). In both panels, the thick black solid line shows the asymptotic logarithmic fit to the direct method measurements (see Equation 2.5). Panel (a): the colored lines represent various strong line calibrations (Tremonti et al. 2004; Zaritsky et al. 1994; Kobulnicky & Kewley 2004; Kewley & Dopita 2002; McGaugh 1991; Denicoló et al. 2002; Pettini & Pagel 2004). Panel (b): the solid, dashed, and dotted gray lines show the median, 68% contour, and 95% contour, respectively, of the Tremonti et al. (2004) MZR. The metallicities and fit parameters for the stacks are reported in Tables 3 and 4 of Andrews & Martini (2013), respectively.
Fig. 2.11.— The direct method $M_* - Z - SFR$ relation for the $M_* - SFR$ stacks (circles color-coded by SFR) in the mass–metallicity plane. The thick solid lines color-coded by SFR show the asymptotic logarithmic fits (see Equation 2.5) for the $M_* - SFR$ stacks. The thick black line shows the direct method MZR from Figure 2.10. The solid, dashed, and dotted gray lines show the median, 68% contour, and 95% contour, respectively, of the Tremonti et al. (2004) MZR. The error bars correspond to the mean error for the $M_* - SFR$ stacks of a given SFR. The metallicities and fit parameters for the stacks are given in Tables 3 and 4 of Andrews & Martini (2013), respectively.
\[
\mu_{0.66} = \log(M_\star) - 0.66 \log(\text{SFR})
\]

Fig. 2.12.— The direct method fundamental metallicity relation for the \( M_\star - \text{SFR} \) stacks (circles color-coded by SFR). The coefficient (0.66) on \( \log(\text{SFR}) \) in the abscissa minimizes the scatter in the FMR (see Equation 2.1). The black line shows a linear fit to the data, with a slope of 0.43.
Fig. 2.13.— The $M_\star$–SFR stacks (circles color-coded by SFR) in the mass–metallicity plane with metallicities determined with the Kobulnicky & Kewley (2004) $R_{23}$ calibration (panel a) and the Pettini & Pagel (2004) N2 calibration (panel b). The thick black line shows the direct method MZR from Figure 2.10. The solid, dashed, and dotted gray lines show the median, 68% contour, and 95% contour, respectively, of the Tremonti et al. (2004) MZR. The error bars correspond to the mean error for the $M_\star$–SFR stacks of a given SFR.
Fig. 2.14.— $N^+$/O$^+$ ratio as a function of direct method oxygen abundance (panel a) and $M_\star$ (panel b) for the $M_\star$ stacks (open circles) and $M_\star$–SFR stacks (circles color-coded by SFR). The horizontal lines show the median of the low oxygen abundance ($12 + \log(O/H) < 8.5$) and low stellar mass ($\log[M_\star] < 8.9$) data. The positively sloped lines in panels (a) and (b) are linear fits to the stacks with $12 + \log(O/H) > 8.5$ and $\log(M_\star) > 8.9$, respectively, whose fit parameters are given in Table 6 of Andrews & Martini (2013). The error bars show the mean error for the $M_\star$ stacks (black) and each SFR bin of the $M_\star$–SFR stacks (color-coded by SFR). If $N^+$/O$^+$ is assumed to trace N/O, as is often done, then our results can be compared directly to literature results on N/O. The N/O abundances of the stacks are reported in Table 3 of Andrews & Martini (2013).
Fig. 2.15.—Our direct method MZR (open circles and thick black line) and the Liang et al. (2007) direct method MZR (blue crosses and line). For reference, the solid, dashed, and dotted gray lines show the median, 68% contour, and 95% contour, respectively, of the Tremonti et al. (2004) MZR.
Fig. 2.16.— Panel (a) shows the excitation parameter, $P = [\text{O} \text{ iii}] \lambda\lambda 4959, 5007 / ([\text{O} \text{ ii}] \lambda 3727 + [\text{O} \text{ iii}] \lambda\lambda 4959, 5007)$, as a function of $R_{23} = ([\text{O} \text{ ii}] \lambda 3727 + [\text{O} \text{ iii}] \lambda\lambda 4959, 5007) / H\beta$. Panel (b) shows $\log([\text{O} \text{ ii}] \lambda 3727 / H\beta)$ versus $\log([\text{O} \text{ iii}] \lambda 5007 / H\beta)$. The gray scale contours (50%, 75%, and 95%) and gray points correspond to SDSS star-forming galaxies. The white and colored circles represent the $M_*$ and $M_*-\text{SFR}$ (color-coded by SFR) stacks, respectively. The light blue contours (50% and 75%) and light blue crosses show $H\text{ ii}$ regions with direct method metallicities from the Pilyugin et al. (2012) compilation. In panel (b), the dashed and dotted lines show lines of constant $[\text{O} \text{ ii}] \lambda 3727 / [\text{O} \text{ iii}] \lambda 5007$ and $R_{23}$, respectively. The stacks trace the overall galaxy distribution better than $H\text{ ii}$ regions, especially at lower excitation parameters. The $H\text{ ii}$ regions tend to have high excitation parameters because the auroral line flux is a strong function of metallicity and hence $R_{23}$. 

88
Chapter 3: Principal Component Abundance Analysis of Microlensed Bulge Dwarf and Subgiant Stars

3.1. Introduction

Elemental abundance trends of Galactic bulge stars are crucial for understanding bulge formation. Galactic chemical evolution is best traced by dwarf stars because their spectra are straightforward to analyze (Edvardsson et al. 1993). However, observations of bulge dwarfs are challenging due to their faintness ($V=19-20$; Feltzing & Gilmore 2000), impeding spectroscopic observation under normal circumstances. Consequently, many studies have focused on giant stars, despite the difficulty in analyzing their spectra. This difficulty has led to shifts in the mean $[\text{Fe/H}]$ and $[\text{Fe/H}]$ distribution function of the bulge as the analysis techniques are refined (e.g., Rich 1988; McWilliam & Rich 1994; Fulbright et al. 2007; Zoccali et al. 2008; Hill et al. 2011). Fortunately, gravitational microlensing offers a unique opportunity to observe bulge dwarfs. When a bulge dwarf is lensed by a foreground object, its brightness can increase by $>5$ magnitudes, enabling spectroscopic observations of sufficiently high resolution and signal-to-noise ($S/N$) for an abundance analysis (e.g., Minniti et al. 1998; Johnson et al. 2007; Bensby et al. 2011 and references therein).
The abundances of different elements are correlated, reflecting their origin in a common nucleosynthetic process, such as Type II or Type Ia supernovae (SNe). However, these correlations are not perfect because the same element can be made in multiple nucleosynthetic processes, whose relative contributions to a star are best distinguished if abundances are measured for many elements. Here we analyze a sample of 35 bulge dwarfs, all of which have at least seven measured elemental abundances and the majority (24/35) of which have 11 or 12 measured elemental abundances, with median errors $\sigma < 0.25$ dex. Bensby et al. (2011) find that the bulge dwarf [Fe/H] distribution function is bimodal, peaked at [Fe/H] $\approx -0.6$ and $+0.3$. They further show that the $\alpha$-element abundances in the bulge dwarfs vary systematically with [Fe/H], following the trends found for thin and thick disk dwarfs (Bensby et al. 2003, 2005; Reddy et al. 2003, 2006). Here we revisit this data set with a different analysis technique based on principal component decomposition of the elemental abundance patterns, showing that the bimodality seen in [Fe/H] also appears in the relative elemental abundance patterns.

Principal component analysis (PCA) is a natural tool for characterizing correlations in a high-dimensional space, reducing the overall dimensionality of the data set while allowing the data themselves to reveal the strongest patterns of correlations. While PCA has a long history in astronomy, its application to elemental

---

1Although our sample includes some subgiant stars, we will describe it as “bulge dwarfs” for brevity.
abundance analysis is relatively new. The only such application we know of is the study of Ting et al. (2012), who used this technique to investigate the distributions in elemental abundance space (hereafter, $C$-space; Freeman & Bland-Hawthorn 2002) defined by various samples of stars from the disk, halo, clusters, and satellite galaxies. They found that disk stars occupied about 6 dimensions within the 17-dimensional $C$-space of the data, but the nucleosynthetic processes likely responsible for the lowest order dimensions changed as a function of [Fe/H]. Their work demonstrated the potential usefulness of principal component abundance analysis (PCAA) as a way to identify groups of stars with distinct enrichment histories. Here we apply this approach to bulge dwarfs to shed further light on the formation of the Galactic bulge.

3.2. Method

PCAA defines a new set of orthogonal basis vectors in $C$-space whose components are chosen to align with the maximum variation within the data not already attributed to lower order components. We use standard PCA (see, e.g., Jolliffe 1986) with the data matrix $\{d_{i,j}\}$ representing the logarithmic abundance relative to iron of element $j$ for star $i$: $d_{i,j} = [X_j/Fe]$ with $X_j = O, Na, Mg, Al, Si, Ca, Ti, Cr, Ni, Zn, Y, and Ba$ for $j = 1, 2, ..., 12$. PCA identifies orthogonal eigenvectors $e_k = \{e_{k,j}\}$

---

\[ \text{Elemental abundances are defined as } [X/Y] \equiv \log(N_X/N_Y) - \log(N_X/N_Y)_\odot, \text{ with missing data replaced by the average value of that elemental abundance for the other stars.} \]
such that the abundance of a given element in a given star can be represented as a sum

\[ d_{i,j} = \bar{d}_j + \sum_{k=1}^{N_{\text{PC}}} c_{i,k} e_{k,j}, \]  

(3.1)

where \( \bar{d}_j = \frac{1}{N_{\text{star}}} \sum_{i=1}^{N_{\text{star}}} d_{i,j} \) is the mean value of \([X_j/Fe]\) in the full data set and \( c_{i,k} \) is the coefficient for the \( k^{\text{th}} \) PC of star \( i \). The first PC describes the direction in elemental abundance space along which the sample stars exhibit the greatest variation, the second PC describes the direction of the second largest variation, etc. If the number of principal components in the sum is equal to the number of elements measured, then the data can be represented exactly. However, if elemental abundances are correlated so that stars are restricted to a lower dimensional subspace, then the elemental abundances can be represented to good accuracy by a smaller number of PCs.

Our data set comes from the homogeneous elemental abundance and error reanalysis of microlensed bulge dwarfs and subgiants by Bensby et al. (2013); abundances for 26 of our 35 sample stars have been previously published (Epstein et al. 2010; Bensby et al. 2010, 2011).

3.3. Results

The dimensionality of the subspace occupied by stars within the full \( C \)-space can be expressed as the number of PCs required to explain the intrinsic variation in the data.
(i.e., the variation not attributable to observational errors). Ting et al. (2012) used Monte Carlo simulations spanning a range of intrinsic dimensionality and variance to show that the true dimensionality was recovered when the cumulative variation of the first \( k \) PCs was about 85\%. For our sample of microlensed bulge dwarfs, the first 1, 2, and 3 PCs describe 64\%, 77\%, and 84\% of the cumulative variation within the data. Although the threshold for completely describing the data likely depends on many factors including sample size, number of observed abundances, and abundance uncertainties, it is fair to say that the bulge dwarfs occupy approximately three dimensions of the 12-dimensional \( C \)-space investigated here.

Figure 3.1 shows the first two PCs derived from the observed bulge dwarf elemental abundances, with the uncertainties determined from bootstrap resampling (see below). PC1 is dominated by the abundances of oxygen, other \( \alpha \)-elements (Mg, Si, Ca, and Ti), and Al, with small uncertainties on the relative contributions of each abundance. SNe II are the primary source of \( \alpha \)-elements and Al, but they are also a significant producer of Fe, especially at early times. By contrast, SNe Ia create large amounts of Fe and other Fe-peak elements, leading to sub-solar \([\alpha/Fe]\) yields; once enough time has elapsed for a substantial number of SNe Ia to occur, they become the dominant source of Fe. The interplay between these two nucleosynthetic sources is thought to underpin the dichotomy in \([\alpha/Fe]\) observed in Galactic disk stars, with high \([\alpha/Fe]\) for “thick disk” stars reflecting rapid formation (e.g., Fuhrmann 1998) and roughly solar \([\alpha/Fe]\) for “thin disk” stars, which have predominantly higher
[Fe/H] (Gilmore & Wyse 1985). Although PCA is a “blind” statistical technique with no \textit{a priori} theoretical input, in this data set (and others we have explored) the first principal component picks up this expected distinction between the two dominant supernova enrichment mechanisms.

PC2 is primarily governed by Na with secondary contributions from Ni (correlated with Na) and Ba (anticorrelated with Na). A similar Na–Ni correlation, attributed to metallicity-dependent SN II yields, has been found previously in halo stars (Nissen & Schuster 1997, 2010); however, this study is the first to identify a Na–Ni correlation amongst bulge stars. Na is primarily produced by hydrostatic carbon burning in the massive stars that explode as SNe II, but proton capture at the same temperatures depletes the pre-explosion Na abundance. As metallicity increases, the neutron excess increases, making Na less susceptible to proton capture and consequently increasing the overall Na yield (Clayton 2003). Similarly, the yield of $^{58}$Ni (the most common Ni isotope) from SNe II is sensitive to the neutron excess and the abundance of neutron-rich nuclei, like $^{23}$Na, in the progenitor star (Woosley et al. 1973). Additional significant $^{58}$Ni production occurs in SNe Ia, whose $^{58}$Ni yield increases with metallicity (Timmes et al. 2003). However, the Na–Ba anticorrelation is not readily explained by a single nucleosynthetic process. One star with distinctive abundances, MOA-2009-BLG-259S (see Figure 3.3c), has a large impact on the contributions of Zn, Y, and Ba to PC2 and drives up the uncertainties for these abundances. The crosses/dashed line in Figure 3.1b show the
effect of omitting this star when defining principal components; PC1 hardly changes, but the contributions of Zn and Y to PC2 switch from being correlated with Na to anticorrelated. The \([\text{Y/Fe}]\) for MOA-2009-BLG-259S has a large uncertainty, though the elevated \([\text{Zn/Fe}] = 0.44 \pm 0.17\) appears to be well-established (Figure 3.3c). Regardless of MOA-2009-BLG-259S, PC2 is dominated by Na and shows a Na–Ni correlation and a Na–Ba anticorrelation.

For comparison, we have found the principal components of a sample of 702 solar neighborhood thin and thick disk dwarfs from Bensby et al. (2003, 2005, and in prep.) with the same elements measured. The disk PC1 and PC2 are shown as triangles/dotted lines in Figure 3.1. The clear similarity between the bulge and disk PC1s suggests that the relative enrichment from SNe II vs. SNe Ia is the main driver of diversity among stars in both samples. On the other hand, the disk and bulge PC2s do not resemble each other: in contrast to the bulge PC2 discussed above, the disk PC2 is dominated entirely by correlated Y and Ba, both neutron-capture elements produced mainly by the \(s\)-process in disk stars (Sneden et al. 2008). The short vs. long lifetimes of the nucleosynthetic sources driving the bulge and disk PC2s, respectively, implies that the bulge formed too rapidly (\(<\approx\)^1 Gyr) for asymptotic giant branch stars to dominate PC2. Johnson et al. (2012) independently reached a similar conclusion based on the relative abundances of \(r\)- and \(s\)-process elements.

Figure 3.2 shows the distribution of the 35 bulge dwarfs in (PC1, PC2)-space, with the upper panel showing the histograms of the bulge dwarfs in PC1 and of
kinematically selected subsets of thin and thick disk dwarfs projected onto the bulge PC1. It is evident from visual inspection that the bulge dwarfs divide into two distinct groups along the PC1 axis, centered at PC1 values of ±0.3. (Because the sum in Equation 3.1 includes the mean elemental abundances of the sample, a value of PC1 = −0.3 corresponds to $\sim [\alpha/Fe]_{\odot}$.) Thus, this analysis of relative elemental abundances, with no direct input from [Fe/H], recovers the bimodality that Bensby et al. (2011) found in the [Fe/H] distribution without reference to relative abundances. Bensby et al. (2011) did find elevated [$\alpha$/Fe] ratios for the metal-poor bulge dwarfs, and we recover the same correlation in this “reverse” analysis: every star with PC1 < −0.1 has [Fe/H] > −0.02, and all but two of the stars with PC1 > −0.1 have [Fe/H] < −0.18. Of these two stars, one (MOA-2009-BLG-259S; square) is a clear PC2 outlier, while the other (MOA-2010-BLG-523S; triangle) is a moderate PC3 outlier that is undistinguished in (PC1, PC2)-space.

Figure 3.2a shows that the metal-rich and metal-poor bulge dwarfs track the thin and thick disk dwarfs, respectively, in PC1. PCAA highlights the scarcity of stars with intermediate [$\alpha$/Fe] (PC1 ~ 0) in the bulge dwarfs and in the thin and thick disk dwarfs. The fact that the two bulge populations track the two disk populations in [Fe/H] and PC1 suggests that the stars had $\alpha$-enrichment histories

---

3The kinematic designation is based on the Bensby et al. (2003) selection criteria; however, we adopt a more stringent cut on the relative thick/thin disk membership probability: stars with $P$(thick disk)/$P$(thin disk) > 100 and $P$(thick disk)/$P$(thin disk) < 0.01 are designated thick and thin disk stars, respectively.
that produce the same abundance patterns. This similarity provides tentative evidence that the bulge formed through secular processes, such as disk instabilities, that heated inner thin and thick disk stars to form the bulge (Kormendy & Kennicutt 2004). Alternatively, the bulge and disk could have formed by distinct mechanisms but experienced “convergent” enrichment histories. For example, Bournaud et al. (2007) propose that the large cosmological accretion rates of high redshift galaxies enable the rapid (∼0.5–1 Gyr), simultaneous formation of the bulge and inner disk, which could account for the $\alpha$-enhanced subpopulation of each component.

To test the statistical significance of PC1 and PC2 and the robustness of the bimodality in Figure 3.2, we created 100 bootstrap resamplings of the data set, redefining principal components each time. PC1 is always recovered, and the general form of PC2 (high Na value, Na–Ni correlation, and Na–Ba anticorrelation) is recovered in 99/100 resamplings. Thus, PC2 is statistically significant even though the contributions of Zn, Y, and Ba to PC2 fluctuate because of its sensitivity to MOA-2009-BLG-259S. The histogram of PC1 values is always multimodal, showing two distinct groupings (as in Figure 3.2a) about 90% of the time; the remaining resamplings show three apparent groupings, but the small sample size makes the distribution of values within the PC1 > 0 group difficult to characterize reliably. The significance of a formal test for bimodality will depend on the form of the adopted null hypothesis, but the likelihood ratio of a 2-Gaussian fit to the PC1 distribution
to a 1-Gaussian fit is $\mathcal{L}(2G)/\mathcal{L}(1G) = 2 \times 10^5$, a large improvement for the addition of two free parameters.

Figure 3.3 shows the decompositions of the elemental abundance patterns of four sample stars into the sum of the sample mean and the first one, two, or three PCs. The best fit coefficients $c_{i,k}$ of Equation (3.1) are found for each star by $\chi^2$-minimization, treating all errors as independent. Panels (a) and (b) show typical examples of metal-poor and metal-rich stars, respectively, each of them fit almost perfectly with a single PC. Panel (c) shows MOA-2009-BLG-259S, which is poorly fit by PC1 alone but well fit when PC2 is also included. Panel (d) shows MOA-2010-BLG-523S, which has one of the largest PC3 coefficients in our sample.

Clearly these PCA fits have low values of $\chi^2_{\text{red}} = \chi^2/(\text{d.o.f.})$, once a sufficient number of PCs (one, two, or three) is included. The number of degrees of freedom is the number of elemental abundances measured for the star minus the number of PCs in the fit. Panel (e) shows the histogram of $\chi^2_{\text{red}}$ values for single-PC and 2-PC decompositions; median values are $\chi^2_{\text{red}} = 0.31$ and 0.21, respectively. The low $\chi^2_{\text{red}}$ values indicate that the observational errors on the elemental abundances are typically overestimated, at least in the sense that they do not represent the variance in estimated elemental abundance that would arise from observing the same star many times. We believe that these low $\chi^2_{\text{red}}$ values arise because the quoted errors effectively incorporate a component of systematic calibration error in addition to random error. For example, multiple lines of the same element may give
discrepant abundance estimates because of uncertainties in the assumed oscillator strengths; the internal dispersion of these estimates is one useful indication of the absolute elemental abundance uncertainty, but the range of elemental abundance measurements from multiple high-S/N spectra will be smaller than this dispersion because the same oscillator strengths are assumed each time. The treatment of random and systematic errors and error correlations is a significant issue for PCAA and other model-fitting approaches, but we defer it to future investigations.

3.4. Conclusions

Our results confirm and extend the findings of Bensby et al. (2011), who identified bimodality in the [Fe/H] distribution of microlensed bulge dwarfs, with metal-poor stars showing enhanced $\alpha$-element abundances like those of solar neighborhood thick disk stars. Our principal component analysis demonstrates that the bimodality found by Bensby et al. (2011) is present even in the relative elemental abundances ([X/Fe]) of the bulge dwarfs alone. The first PC is dominated by $\alpha$-elements, reflecting the dichotomy between SN II and SN Ia enrichment. The second PC captures the Na–Ni correlation caused by the metallicity dependence of SN II yields. Intriguingly, the two metal-rich stars that exhibit $\alpha$-enhancement (MOA-2009-BLG-259S and MOA-2010-BLG-523S) are also outliers from the main locus of stars in (PC1, PC2, PC3)-space, suggesting that they do indeed have unusual enrichment histories. If we project the thin and thick disk dwarfs onto the bulge dwarf PC1, they occupy
the locations of the metal-rich and metal-poor bulge dwarfs, respectively. Analyzing local disk dwarfs, we find that the first principal component is nearly identical to the bulge PC1. However, the disk PC2 is governed by Y and Ba, products of long-lived asymptotic giant branch stars, whereas enrichment from short-lived SNe II drives the bulge PC2. Qualitatively, these results support a scenario in which the bulge grows by secular evolution from the inner disk, which itself has the elemental abundance dichotomy seen in local populations, but there may be other bulge formation models that can produce similar results.

Our results, and those of Ting et al. (2012), illustrate the potential of PCAA as a tool for characterizing the distribution of stars in high-dimensional C-space. One application, highlighted here, is to identify subpopulations in a sample, drawing on the information present in all measured elemental abundances simultaneously. With large multi-element samples, this approach could be used to isolate “interloper” stars accreted from a dissolved satellite, and perhaps to identify cohorts of stars associated with common birth clusters (Freeman & Bland-Hawthorn 2002). A second application, illustrated by the examples of MOA-2009-BLG-259S and MOA-2010-BLG-523S, is to identify outlier stars, either through their unusual locations in PC-space or because they are poorly fit by combinations of PCs that fit most stars well. Such outliers may reveal rare but physically informative enrichment pathways. A third application, highlighted by Ting et al. (2012), is to characterize the dimensionality of the stellar distribution in C-space, which is a basic test for
models of Galactic assembly and enrichment. We will explore this technique in future work using theoretical models.

By allowing the data themselves to define the directions of strongest variation, PCAA complements the usual approach of testing predictive models that adopt nucleosynthetic yields from theoretical calculations. The PCs clearly do have physical content, but the connection of PCs to enrichment mechanisms is not one-to-one (see Ting et al. 2012), and interpreting them will require comparisons to models that vary both enrichment and mixing histories and the nucleosynthetic yields themselves. Our analysis identifies several practical complications of PCAA, including the potential sensitivity to outliers in small samples, the treatment of missing elemental abundance measurements, the impact of heteroscedastic errors and correlated errors, and the mix of random and systematic contributions to the errors quoted in observational analyses. We will investigate these issues in future work.

The enormous samples of high-resolution spectra anticipated from the SDSS-III APOGEE survey (Eisenstein et al. 2011, Majewski et al. in prep.), the Gaia-ESO survey (Gilmore et al. 2012), and the GALAH survey (Barden et al. 2010) will map the elemental abundance distribution over wide swaths of the Galaxy, and PCAA will be a valuable tool for connecting these measurements to a comprehensive theory of the formation of the Milky Way.
3.5. Figures
Fig. 3.1.— A graphical representation of the PC1 (panel a) and PC2 (panel b) vectors of the microlensed bulge dwarfs (circles/solid lines) and the disk dwarf sample of Bensby et al. (2003, 2005, and in prep.) (triangles/dotted lines). The PC1 and PC2 eigenvector components refer to $e_{1,j}$ and $e_{2,j}$, respectively, from Equation 3.1, with uncertainties from bootstrap resampling. We also show PC2 for the bulge dwarfs if MOA-2009-BLG-259S is omitted from the sample (crosses/dashed line in panel b); PC1 remains nearly unchanged, so it is not plotted for clarity. Each symbol represents an abundance ([X/Fe]) and thus a dimension in the original C-space.
Fig. 3.2.— Panel (a): the histograms of bulge dwarfs (gray) and of kinematically selected thin (red) and thick (blue) disk dwarfs projected onto the bulge PC1 axis defined above. Panel (b): the distribution of bulge dwarfs in (PC1, PC2)-space, color-coded by [Fe/H]. The square and triangle represent the PC2 outlier (MOA-2009-BLG-259S) and the PC3 outlier (MOA-2010-BLG-523S), respectively. The positions along PC1 and PC2 correspond to $c_{i,1}$ and $c_{i,2}$, respectively, in Equation 3.1.
Fig. 3.3.— Panels (a)–(d) show elemental abundances for, respectively, a metal-poor star, a metal-rich star, the PC2 outlier star (MOA-2009-BLG-259S), and the PC3 outlier star (MOA-2010-BLG-523S). Lines indicate the best fit combined abundance pattern using the mean abundance and the first PC (solid gray line), the first two PCs (short dashed line), and the first three PCs (long dashed line in panel (d) only). Reduced-$\chi^2$ ($\chi^2_{\text{red}}$) values for these fits are listed in the legends. Panel (e) shows the distribution of $\chi^2_{\text{red}}$ values for the bulge dwarf sample, with the same line types as panels (a)–(c).
4.1. Introduction

Classical chemical evolution models are powerful tools for testing theories of galaxy evolution because of the tight constraints imposed by stellar elemental abundances. Early work by Schmidt (1959, 1963) introduced a basic one-zone model of chemical evolution. Subsequent investigations have become more intricate with time (e.g., Truran & Cameron 1971; Tinsley 1974, 1976, 1977; Matteucci & Francois 1989; Chiappini et al. 1997, 2001; Schönrich & Binney 2009a) to reproduce improving observational constraints. However, the complexity of more recent models can obscure the effects of and trade-offs between model parameters and assumptions, which are best highlighted with one-zone models. An interesting limiting case with an analytic solution is the so-called “Simple” model (Schmidt 1959, 1963). The Simple model assumes a one-zone closed system of initially pristine gas, a constant stellar initial mass function, and complete and instantaneous mixing of the gas reservoir. While the Simple model does not reproduce all observational constraints, such as the metallicity distribution function of the solar neighborhood (known as...
the G-dwarf problem; van den Bergh 1962; Schmidt 1963), it is still frequently used as an illustrative example. The Simple model can be further idealized by adopting the instantaneous recycling approximation (Talbot & Arnett 1971; Searle & Sargent 1972), which treats stars that are more massive than the Sun as evolving instantaneously and stars that are less massive than the Sun as living forever. This approximation works well for primary elements, which are the synthesized from hydrogen and helium, but not for secondary elements, whose yields depend on the initial metal content of the star. Since our goal is to isolate the effects of and illustrate the interactions between the ingredients of chemical models in a realistic manner, we developed a more advanced but still straightforward one-zone model that is an open system (i.e., with inflow and outflow) and relaxes the instantaneous recycling approximation.

4.2. Model Description

We have constructed a flexible one zone chemical evolution model with inflow and outflow (i.e., open box) whose gas is instantaneously and completely mixed. The model is evolved for 12 Gyr with time steps every 30 Myr, which corresponds to the lifetime of the longest-lived star that will explode as a CCSN. The model includes enrichment from CCSNe, SNeIa, and AGB stars, and the yields of the latter two sources are staggered (i.e., not instantaneously recycled). The model stochastically draws individual stars from the stellar initial mass function (IMF) and stochastically
explodes white dwarfs (WDs) as SNeIa. The stochastic nature of the model enables investigations of scatter in regimes where the IMF is not fully populated. As the total mass increases and the IMF becomes well-sampled, the results from our numerical approach quickly converge to the those of analytic calculations.

4.2.1. Fiducial Simulation

The parameters of the fiducial simulation were chosen to broadly match the observed trend in \([O/Fe]–[Fe/H]\). It is a one zone open box with inflow and outflow. We adopt the CCSN yields from Chieffi & Limongi (2004) and Limongi & Chieffi (2006), AGB star yields from Karakas (2010), and the W70 SNIa yields from Iwamoto et al. (1999) (see Section 4.4.2). The simulation starts with \(2 \times 10^{10} \, M_\odot\) of primordial gas and additional primordial gas flows into the galaxy with an exponentially declining time profile \((\tau = 6 \, \text{Gyr})\). The final total mass of the simulation is \(6.9 \times 10^{10} \, M_\odot\), with \(3.2 \times 10^9 \, M_\odot\) in gas, \(5.2 \times 10^{10} \, M_\odot\) in stars, and \(1.4 \times 10^{10} \, M_\odot\) in remnants (black holes, neutron stars, and white dwarfs). The star formation rate (SFR) is set by a constant star formation efficiency (SFE) of \(10^{-9} \, \text{yr}^{-1}\), converting \(M_{\text{gas}}(t) \cdot (\Delta t/1 \, \text{Gyr})\) of the available gas mass into stars in a time interval \(\Delta t\). The cold ISM is ejected in an outflow at a rate of \(2.5 \times \text{SFR}\). We adopt a Kroupa (2001) IMF from 0.1–100 \(M_\odot\) and the stellar lifetime function of Padovani & Matteucci (1993). The SNIa delay time distribution (DTD) is an exponential with a timescale of 1.5 Gyr and
a minimum delay time of 150 Myr. The parameters of the fiducial simulation are explained in more detail in Section 4.3.

4.2.2. Closed Box and Inflow-only Simulations

To illustrate the importance of accretion and outflows, Figure 4.1 compares the evolution of the fiducial simulation to that of a closed box (no gas inflow or outflow) simulation and an inflow-only simulation. The closed box simulation reaches a very high metallicity ([Fe/H] ≈ 1.5) because most of the gas has been turned into stars ($M_{\text{gas}}^{\text{final}} = 2.0 \times 10^8 \, M_\odot$), leaving only a small amount of hydrogen to dilute the iron production from CCSNe and SNeIa. The inflow-only simulation achieves a lower but still excessive metallicity of [Fe/H] ≈ 0.9 and finishes at a higher [O/Fe] (0.05 vs. −0.05 for the closed box simulation), because relative to a closed box the inflow dilutes the metallicity by providing additional hydrogen and fuels more star formation and CCSN enrichment at late times. Comparison to the fiducial model demonstrates the effect of outflows, which decrease the overall metallicity ([Fe/H] ≈ 0 at the end of the simulation by design) by removing metals and the gas that could have fueled more star formation and metal production. Outflows have the secondary effect of decreasing [O/Fe] because the enrichment products created earlier, like oxygen, have more opportunities to be ejected than enrichment products returned on longer timescales, like iron. The fiducial simulation roughly reproduces the observed abundance trend in the data of Ramírez et al. (2013), though we
caution that the detailed appearance of the observations depends on sample selection and on the methodology for abundance estimates. In the fiducial model and all of our other simulations that approximately reproduce observed abundance trends, the abundance evolution slows dramatically at late times, approaching an approximate equilibrium between sources and sinks of different elements. Below we refer to these approximately asymptotic states as equilibrium abundances.

In Figure 4.2 we show the mass of gas-phase oxygen and iron as a function of metallicity (stellar abundances are taken to be the gas-phase abundances when a stellar generation is born). For ease of interpretation we show the results from a constant SFR simulation because its iron mass is linear with [Fe/H]; this simulation follows a very similar locus in [O/Fe]–[Fe/H] to the fiducial simulation, though its final [Fe/H] is lower and [O/Fe] is higher (see Figure 4.3b). The solid line shows the total gas-phase mass of each element. The dotted, short-dashed and long-dashed lines indicate the contributions from CCSNe, AGB stars, and SNeIa, respectively. The ordinate axes have been offset such that the total oxygen and iron lines cross at solar [O/Fe].

One can metaphorically regard CCSNe, SNeIa, and AGB stars as tributaries that enrich the “river” of the main gas reservoir, which is itself depleted by star formation and outflow. Far “upstream,” at early times and low [Fe/H], CCSNe are the only significant channel and the ratio [O/Fe] ≈ 0.4 reflects their high-α yields. As SNeIa become important, contributing iron but minimal oxygen, the [O/Fe] ratio
declines toward solar values. The transition from CCSNe to SNeIa dominating iron production occurs at $t = 2.7$ Gyr and $[\text{Fe/H}] \approx -0.17$. CCSNe remain the dominant source of oxygen production at all times, with a much smaller contribution from AGB stars. Since these lines include recycled material, AGB stars “contribute” iron even though they do not directly synthesize it because they return iron incorporated at birth. For similar plots of the full suite of elements and more in-depth discussion of multi-element abundances see Figure 4.12 and Section 4.4.2.

4.3. Varying Parameters: Model Tracks in $[\text{O/Fe}]$ vs. $[\text{Fe/H}]$

In this section, we vary the parameters of the model to understand how galaxy evolution parameters affect the mean track in $[\text{O/Fe}]$ vs. $[\text{Fe/H}]$. The mean track in this space reflects the relative enrichment from CCSNe and SNeIa ($[\text{O/Fe}]$) as a function of overall metallicity ($[\text{Fe/H}]$). Since metallicity maps onto time nearly monotonically in the simulations, the mean track also encodes the history of CCSNe versus SNeIa enrichment.

4.3.1. Inflow Rate

Closed box chemical evolution models generically suffer from the G-dwarf problem (van den Bergh 1962), namely that they produce metallicity distribution functions for G-dwarfs, whose lifetimes are comparable to the age of the Galactic disk, with too many metal-poor stars relative to metal-rich stars. Inflows allow galaxies to form
more stars later in their lifetimes when the cold ISM is metal-rich, increasing the
fraction of metal-rich stars. Continuing infall is, of course, the natural expectation in
analytic or numerical models of cosmological galaxy formation, with star formation
typically tracking gas accretion after a moderate delay (e.g., Katz et al. 1996).

To represent a range of inflow histories, we vary the timescale $\tau_1$ assuming an
exponential inflow rate:

$$\dot{M}_{\text{in}} = \frac{M_1}{\tau_1} e^{-t/\tau_1},$$  \hspace{1cm} (4.1)

where $M_1$ sets the normalization of the inflow rate. Figure 4.3a shows the effect of
varying the inflow timescale by a factor of two relative to the $\tau_1 = 6$ Gyr of the
fiducial simulation. The 3 Gyr simulation enriches more quickly, and it reaches a
higher [Fe/H] at the “knee” of the enrichment curve because it has more CCSNe
before SNeIa start to drive [O/Fe] down. Its equilibrium [O/Fe] is higher and [Fe/H]
lower than those of the fiducial simulation due to less late time accretion and star
formation, though it follows the same track as the fiducial simulation after $t = 1$ Gyr.
Conversely, the longer inflow timescale ($\tau_1 = 12$ Gyr) simulation has a knee at a
marginally lower [Fe/H] but reaches a lower [O/Fe] and higher [Fe/H] at the end of
the simulation. Its track also overlaps the fiducial track. Overall, inflow timescale
slightly alters the location of the [O/Fe] knee, but it mostly dictates how far along
the track the simulation goes until it reaches equilibrium abundance ratios at late
times.
We experimented with several functional forms of the inflow history with different ratios of early to late time accretion, including a double exponential

$$\dot{M}_{\text{in}} = \frac{M_1}{\tau_1} e^{-t/\tau_1} + \frac{M_2}{\tau_2} e^{-t/\tau_2},$$

and a linear-exponential product

$$\dot{M}_{\text{in}} = \left(\frac{M_1}{\tau_1}\right) \left(\frac{t}{\tau_1}\right) e^{-t/\tau_1}.$$  

(4.2)

(4.3)

We also ran a constant SFR simulation, where the inflow rate was set to maintain a constant gas mass at all times. This simplified star formation history (SFH) helps illustrate the differences between simulations with different yields (see Figure 4.11) and the relative contribution of CCSNe, SNeIa, and AGB stars to enrichment (see Figures 4.2 and 4.12).

Figure 4.6 shows the SFR as a function of time for the fiducial, double exponential, linear-exponential, and the constant SFR simulations. The double exponential simulation represents a scenario with early rapid accretion ($M_1 = 10^{10} M_\odot$, $\tau_1 = 300$ Myr) followed by steady accretion on long timescales ($M_2 = 6 \times 10^{11} M_\odot$, $\tau_2 = 14$ Gyr), where the timescales were adopted from Schönherr & Binney (2009a). The linear-exponential simulation is motivated by cosmological hydrodynamical simulations that find inflow rates that increase at early times, reach a maximum at a characteristic timescale, and then decline at late times (Simha et al. 2014). The timescale of the linear-exponential simulation ($\tau_1 = 3.5$ Gyr) was chosen so
that the simulation reached solar metallicity and oxygen abundance. In contrast
to the declining SFHs of the other simulations, the constant SFR simulation
maintains a high SFR at late times. The normalizations of the double exponential,
linear-exponential, and the constant SFR simulations were chosen to approximately
match the same total amount of inflow and final stellar mass as the fiducial
simulation.

Figure 4.3b shows the mean tracks in \([\text{O/Fe}]\) vs. \([\text{Fe/H}]\) for these various inflow
histories. The double exponential simulation enriches slightly faster than the fiducial
simulation for the first 2.5 Gyr, but its higher late time inflow rate results in a
lower equilibrium \([\text{Fe/H}]\) and a higher equilibrium \([\text{O/Fe}]\). The linear-exponential
simulation enriches rapidly in the first 500 Myr because the small amount of
pristine gas that is accreted does little to dilute the enrichment from CCSNe
and SNeIa that formed from the initial gas mass. However, as the inflow rate
increases \((t = 500 \text{ Myr} - 2 \text{ Gyr})\), the track moves downward due to SNIa enrichment
plus dilution from the pristine inflow. When the inflow rate starts to decline,
the track overlaps the fiducial simulation. As mentioned previously, the constant
SFR simulation follows nearly the same track as the fiducial simulation, but its
equilibrium \([\text{Fe/H}]\) is lower and a equilibrium \([\text{O/Fe}]\) is higher due to the additional
dilution and CCSN enrichment at late times from its higher late inflow rate.

The total mass of the simulation does not affect the mean trend in \([\text{O/Fe}] -
\[\text{Fe/H}\]\), as long as the initial mass and the mass inflow rate are scaled together.
However, larger masses decrease the effect of stochastic fluctuations from incomplete sampling of the IMF or the SNIa DTD. For a fixed inflow rate, the initial gas mass of the model regulates the speed of enrichment but not the final metallicity. A smaller initial gas mass extends the plateau in \([O/Fe]-[Fe/H]\) to lower metallicities. A large initial gas mass produces a galaxy that enriches quickly then asymptotes to a constant metallicity, whereas a small initial gas mass results in a smoother, more continual increase in metallicity.

### 4.3.2. Star Formation Rate and Efficiency

We assume that a constant fraction of the gas turns into stars in each time step:

\[
\text{SFR} = \frac{1}{t_{\text{gas}}} M_{\text{gas}} = \text{SFE} \cdot M_{\text{gas}},
\]

where \(t_{\text{gas}}\) is the gas depletion timescale and \(\text{SFE} \equiv t_{\text{gas}}^{-1}\) is the star formation efficiency. Figure 4.4a shows factor of three variations around the SFE of the fiducial simulation (\(10^{-9} \text{ yr}^{-1}\)). Increasing the SFE results in a knee at a higher \([Fe/H]\) because more CCSNe produce more iron and more star formation consumes more hydrogen. Further increases in the SFE result in diminishing increases in the \([Fe/H]\) of the knee because the simulation reaches a temporary equilibrium abundance. The final equilibrium \([Fe/H]\) and \([O/Fe]\) are insensitive to SFE, though SFE sets the speed of convergence.
4.3.3. Outflows

The energy and momentum injection from massive stars can launch large-scale galactic winds (e.g., Veilleux et al. 2005). The mass outflow rates of these winds are not well-constrained observationally, but hydrodynamic galaxy formation simulations find that outflow rates of $1–10 \times \text{SFR}$ are required to reproduce observed galaxy stellar masses and morphologies (e.g., Hopkins et al. 2012). We parameterize the outflow rate ($\dot{M}_{\text{outflow}}$) to be proportional to the SFR,

$$\dot{M}_{\text{outflow}} = \eta_{\text{wind}} \text{SFR},$$

where $\eta_{\text{wind}}$ is the mass-loading factor of the wind. We assume that the outflowing material has the same abundance pattern as the cold ISM, which would be expected if the mass in the outflow is dominated by entrained cold gas as opposed to the hot gas ejected by SNe. We performed simulations that ejected a fraction of the stellar yields without any cold ISM leaving the Galaxy, which required 75% of the stellar yields to be ejected for the simulation to have an equilibrium $[\text{O/Fe}]$ and $[\text{Fe/H}]$ at solar values but the knee occurred at $[\text{Fe/H}] = -1.5$.

Figure 4.4b shows the mean track in $[\text{O/Fe}]-[\text{Fe/H}]$ for factor of two variations in the mass-loading factor around the fiducial value of $\eta_{\text{wind}} = 2.5$, which produces a track that ends at solar metallicity and oxygen abundance. Changing $\eta_{\text{wind}}$ strongly affects the shape of the track after the knee at $[\text{Fe/H}] \approx -0.9$, leading to very
different equilibrium abundances. Increasing the mass-loading factor removes more metal-rich gas, which is replaced by inflowing pristine gas, thus decreasing the equilibrium metallicity. Elements produced earlier in the simulation (like oxygen) have more opportunities to get ejected than those that are produced later (such as iron), which causes the decrease in the equilibrium \([O/Fe]\) value at high \(\eta_{\text{wind}}\).

A higher mass-loading factor also induces greater variance in the abundances because it decreases the size of the gas reservoir and boosts the impact of stochastic fluctuations, but this effect only becomes important at much lower total masses than shown in Figure 4.4b.

4.3.4. Inflow Metallicity and Metal Recycling

Gas that accretes onto galaxies is a mix of pristine or low metallicity \((Z = 10^{-3} \, Z_\odot)\); Songaila & Cowie (1996) IGM gas and enriched material previously ejected from galaxies (e.g., Ford et al. 2013). The composition of accreting gas is difficult to determine observationally, so the relative contribution of new vs. recycled gas is unknown. We performed simulations where the overall metallicity of the inflowing gas was

- \(Z = 0\) (fiducial value used to simulate inflow of pristine gas),

- \(Z = 10^{-2} \, Z_\odot\) (to simulate inflow from an enriched IGM), and

- \(Z = 10^{-1} \, Z_\odot\) (to simulate a mix of pristine gas and recycled gas),
where $Z_\odot = 0.02$. The composition of the inflow was the relative abundances of elements in the ISM at the previous time step, which was then scaled to an overall metallicity. We also experimented with $\alpha$-enhanced and scaled-solar abundance patterns but chose not to use them because the $\alpha$-enhanced abundance pattern prevented the gas from reaching the solar [$\alpha$/Fe] value and the scaled-solar abundance pattern created very low metallicity stars with solar [$\alpha$/Fe] abundances.

Figure 4.5a shows the simulations with enriched inflow. The $Z = 10^{-2} Z_\odot$ simulation is nearly identical to the fiducial ($Z = 0$) simulation, indicating that enriched inflow from the IGM has little effect on chemical evolution (though it may be important at extremely low metallicities; see Brook et al. 2013). The $Z = 10^{-1} Z_\odot$ simulation has a higher [Fe/H] and a higher [O/Fe] (above the knee) at fixed time than the fiducial simulation, so the knee occurs at a higher [Fe/H]. We conclude that enriched inflow only has a major effect on the track in [O/Fe] vs. [Fe/H] if the accreted gas has a metallicity of $Z > 10^{-1} Z_\odot$.

4.3.5. IMF

The slope and mass range of the stellar initial mass function (IMF) sets the relative number of high and low mass stars that form in a stellar population. The IMF affects the production of elements whose CCSNe yields are highly mass dependent, like oxygen. It also determines the ratio of CCSNe to SNeIa (at fixed SNIa fraction), which governs the timescale for iron enrichment.
Historically, the Salpeter (1955) IMF, with a slope of $\alpha = 2.35$ and a mass range from 0.1–100 $M_\odot$, has provided a useful reference point for IMFs. However, the more recent Kroupa (2001) IMF, whose slope is $\alpha = (1.3, 2.3)$ from (0.1–0.5, 0.5–100 $M_\odot$) does a better job describing the solar neighborhood, so we adopt it as the fiducial IMF.

In Figure 4.5b, we show the mean track in [O/Fe]–[Fe/H] for four IMFs:

- Kroupa IMF from 0.1–100 $M_\odot$ (fiducial),
- flat ($\alpha = 2.1$) IMF from 0.1–100 $M_\odot$,
- Kroupa IMF from 0.1–50 $M_\odot$, and
- Salpeter IMF from 0.1–100 $M_\odot$.

Relative to the fiducial IMF, the flat IMF simulation always has a higher [O/Fe]. It also has a slightly higher [Fe/H], especially at early times ($t < 2$ Gyr). The Kroupa 0.1–50 $M_\odot$ IMF simulation has a lower [O/Fe] but a similar [Fe/H] compared to the fiducial simulation. These changes are due to the oxygen yield of CCSNe increasing strongly with mass but iron yield staying constant.

The Salpeter 0.1–100 $M_\odot$ IMF simulation has a nearly identical [O/Fe] value of the plateau because of the similar high mass slopes of the Kroupa and Salpeter IMFs and the lack of SNeIa iron production at these early time steps. However, the
larger number of low mass stars that are produced in the Salpeter IMF simulation results in a lower [Fe/H] because more gas goes into long-lived low mass stars as opposed to CCSNe.

Although Figure 4.5b shows that chemical evolution tracks are sensitive diagnostics of the IMF, these differences are partly degenerate with SN yields. The mass cut for CCSNe, which sets the division between mass ejected in the explosion and mass that falls back on the neutron star, has a large impact. In particular, the iron yield is governed by the mass cut because the material near the mass cut consists mostly of iron-peak elements. Decreasing the mass cut to produce 0.05 M⊙ of ejected 56Ni instead of 0.1 M⊙, increases the [O/Fe] level of the plateau, decreases the [Fe/H] of the knee, and generally decreases [Fe/H] at all times. We discuss mass cut effects further in Section 4.4.1

4.3.6. SNIa Delay Time Distribution

The SNIa history cannot be predicted robustly due to uncertainty in the delay time between the birth of a stellar population and the SNeIa that it will produce. Our understanding of the SNIa delay time distribution (DTD) is complicated by a few factors: (1) the evolutionary state of the progenitor systems—double degenerate (WD + WD) or single degenerate (WD + red giant, WD + main sequence star, or WD + helium star)—remains unknown, (2) binary evolution, especially the common envelope phase, is difficult to model, and (3) dynamical effects in triple star systems
could dramatically hasten the merger of two WDs relative to naive expectations from binary evolution (Thompson 2011).

Because of these uncertainties, we consider variations in the minimum delay time and multiple functional forms of the DTD. The minimum delay time sets the timescale for SNeIa to become significant contributors of iron, which produces the knee in [O/Fe]–[Fe/H]. The different functional forms of the DTD govern the SNIa rate on longer timescales.

In the fiducial simulation, we adopt an exponential SNIa DTD as in Schönrich & Binney (2009a), which they chose to match the decline in [O/Fe] and [Ca/Fe] data. The total mass of WDs that will explode as SNeIa ($M_{\text{SN} \text{ Ia}}^{\text{WD}}$) is set to be 13.5% of the mass in WDs formed by stars of initial mass between 3.2–8.0 $M_\odot$. This value was chosen to match the observed time-integrated (from 40 Myr to 10 Gyr) SNIa rate of $N_{\text{Ia}}/M_\star = 2.2 \times 10^{-3} M_\odot^{-1}$ from Maoz & Mannucci (2012) (their adopted Bell et al. 2003 “diet Salpeter” IMF produces a nearly identical amount of mass as the Kroupa IMF). For comparison, Schönrich & Binney (2009a) assumed that 7.5% of the mass in WDs formed from stars of initial mass between 3.2–8.5 $M_\odot$ exploded as SNeIa for a Salpeter IMF, which corresponds to 5.5% of the mass in WDs formed from stars of initial mass between 3.2–8.0 $M_\odot$ for a Kroupa (2001) IMF. The reservoir of
WDs that will explode as SNeIa is depleted according to Equation (5) of Schönrich & Binney (2009a):

\[
\frac{dM_{WD}^{SNIa}}{dt} = 0 \text{ for } t < t_{SNIa}^{\text{min}}
\]

and

\[
\frac{dM_{WD}^{SNIa}}{dt} = -\frac{M_{WD}^{SNIa}}{\tau_{SNIa}} \text{ for } t > t_{SNIa}^{\text{min}},
\]

where \(t_{SNIa}^{\text{min}}\) is the minimum delay time for a SNIa and \(\tau_{SNIa}\) is the timescale for SNeIa, which are set to 150 Myr and 1.5 Gyr from the birth of a stellar population, respectively, in the fiducial sim. Equivalently, the SNIa DTD is

\[
f_{SNIa}(t) = \frac{m_{WD}^{SNIa} \tau_{SNIa}^{-1} e^{-(t-t_{SNIa}^{\text{min}})/\tau_{SNIa}}}{m_{SNIa}},
\]

where the units of \(f_{SNIa}(t)\) are SNIa yr\(^{-1}\) M\(_\odot\)^{-1}, \(m_{WD}^{SNIa}\) is the mass fraction of a stellar population that will go into WDs that explode as SNeIa, and \(m_{SNIa}\) is the ejecta mass of an SNIa (1.37 M\(_\odot\) in the W70 model of Iwamoto et al. 1999). We have assumed an approximate relationship between WD mass and initial mass of \(m_{WD} = 0.085 m_{\text{initial}} + 0.48\) based on the Karakas (2010) AGB yields.

Figure 4.7 shows the mean tracks in [O/Fe]–[Fe/H] for an exponential SNIa DTD with three different minimum delay times: 40, 150 (fiducial value), and 300 Myr. The fiducial simulation turns over at [Fe/H] \sim -0.9. Increasing the
minimum delay time increases the [Fe/H] of the knee and makes the turnover sharper. After about 2 Gyr, the [O/Fe]–[Fe/H] tracks of the three simulations are very similar. The data do not rule out any of these possible delay times.

We tested four functional forms of the SNIa DTD: an exponential DTD (fiducial functional form), a theoretical DTD based on the single degenerate scenario, an empirical power law DTD, and an empirical prompt + delayed DTD. Figure 4.8a shows the DTD, represented by \( f_{\text{SNIa}}(t) \), for the first three functional forms. (The prompt + delayed DTD depends on the SFH, so it cannot be plotted for a single stellar population.) These DTDs were normalized to match the observed time-integrated (from 40 Myr to 10 Gyr) SNIa rate of \( N_{\text{Ia}}/M_\star = 2.2 \times 10^{-3} M_\odot^{-1} \) from Maoz & Mannucci (2012). Figure 4.8b shows the number of SNeIa that explode as a function of time for each of the four DTDs. For the prompt + delayed model we adopted the normalization of Scannapieco & Bildsten (2005), but for the other three DTDs we normalized to \( N_{\text{Ia}}/M_\star = 2.2 \times 10^{-3} M_\odot^{-1} \).

We implement the theoretical DTD from Greggio (2005) for single degenerate systems that reach the Chandrasekhar limit. This scenario (described in Section 3.1 of Greggio 2005) requires the primary star to form a carbon–oxygen (CO) WD (i.e., \( m_{\text{WD}} \geq 0.6 \, M_\odot \)), which corresponds to an initial mass of \( m_1 \geq 2 \, M_\odot \) (first break in the single degenerate curve in Figure 4.8a). In addition, the WD must be able to accrete enough material from the secondary to reach the Chandrasekhar mass, which depends on the envelope mass of the secondary (itself a function of the initial mass.
of the secondary) multiplied by the accretion efficiency. Thus, systems with low mass secondaries require larger WDs and hence higher mass primaries to achieve the Chandrasekhar mass (second break in the single degenerate curve in Figure 4.8).

Recent measurements of the SNIa rates and the SFHs of galaxies have placed interesting observational constraints on the SNIa DTD (Totani et al. 2008; Maoz & Badenes 2010; Maoz et al. 2010, 2011, 2012; Graur et al. 2011; Maoz & Mannucci 2012). These studies face a trade-off between (1) observing more SNeIa, which requires a larger volume or more galaxies surveyed, and (2) having a better understanding of the star formation histories of the galaxies surveyed, which requires resolved stellar populations, spectral-synthesis-based SFHs, galaxy colors, or the cosmic SFH (see Maoz & Mannucci 2012 for a detailed description of particular techniques). Encouragingly, the results from different techniques seem to agree that a $t^{-1}$ power law DTD provides a good fit to the data,

$$f_{\text{SNIa}}(t) = a \left( \frac{t}{\text{yr}} \right)^{-1}, \quad (4.9)$$

where $f_{\text{SNIa}}(t)$ is in units of SNIa yr$^{-1}$ M$_\odot^{-1}$, and we adopt $a = 3.98 \times 10^{-4}$ to recover $N_{\text{Ia}}/M_* = 2.2 \times 10^{-3}$ M$_\odot$ from Maoz & Mannucci (2012). The evidence for a $t^{-1}$ power law DTD is strongest for $t = 1$–10 Gyr (Maoz & Mannucci 2012), but the data are consistent with a $t^{-1}$ power law continuing to shorter delay times with a minimum delay time of 40 Myr from the birth of a stellar population, which
is approximately the lifetime of the most massive star that will produce a WD as opposed to a neutron star.

We also consider the empirical prompt + delayed DTD based on observations that the SNIa rate is correlated with SFR tracers \( \text{Scannapieco & Bildsten} \) (2005; \text{Manucci et al.} 2006). In this scenario, the “prompt” SNIa are associated with young stellar populations, so their rate scales with the SFR. The “delayed” SNIa come from old stellar populations, so their rate depends on stellar mass. The SNIa rate is

\[
\frac{R_{\text{SNIa}}(t)}{yr^{-1}} = A \left( \frac{M_*}{M_\odot} \right) + B \left( \frac{\text{SFR}}{M_\odot yr^{-1}} \right),
\]

(4.10)

where \( A = 4.4 \times 10^{-14} \) and \( B = 2.6 \times 10^{-3} \) are the values adopted by Scannapieco & Bildsten (2005) converted to the relevant units, and we have assumed \( t_{\text{SNIa min}} = 40 \) Myr. Maoz & Mannucci (2012) point out that this scenario is consistent with coarse time sampling of a \( t^{-1} \) power law DTD, and the demarcation between the prompt and delayed components is somewhat arbitrary.

Figure 4.9 shows the mean tracks in \([O/Fe] – [Fe/H]\) of simulations with various SNIa DTDs. We also include an exponential DTD with \( t_{\text{SNIa min}} = 40 \) Myr (blue dashed line) to more closely match the other DTDs. The power law DTD simulation turns over more dramatically than the exponential DTD, but it reaches the same [Fe/H] and [O/Fe] at about 2 Gyr. It then continues along the same track at a slower pace.
but ends at a higher equilibrium $[\text{Fe/H}]$ and lower $[\text{O/Fe}]$ than the exponential DTD. The single degenerate scenario DTD has a similar track to the power law DTD over the first 500 Myr. It reaches the same track as the exponential DTD but finishes with a lower equilibrium $[\text{Fe/H}]$ and a higher $[\text{O/Fe}]$ (near the 4 Gyr symbol). The exponential, power law, and single degenerate scenario DTD are all consistent with the $[\text{O/Fe}]$–$[\text{Fe/H}]$ data, especially if other parameters, such as the inflow timescale or the SNIa normalization, are allowed to vary to compensate for the different equilibrium $[\text{Fe/H}]$ and $[\text{O/Fe}]$. On the other hand, the prompt + delayed DTD creates a sudden drop in $[\text{O/Fe}]$ before leveling off, a feature that is not present in the data. This behavior is caused by the large number of prompt SNeIa (see Figure 4.8b) and is a generic prediction of the prompt + delayed DTD (see also Figure 5 of Matteucci et al. 2009).

4.3.7. Warm ISM

Stellar ejecta from CCSN, SNIa, and AGB stars may not be immediately returned to the cold ISM. Instead, stellar yields may be injected into the warm ($T > 10^4$ K) ISM that does not form stars. Over time, gas in the warm ISM cools into the cold ISM and can get reincorporated into future generations of stars. While we have not included a warm ISM in the fiducial simulation, we explore its effects on chemical evolution in Figure 4.10.
For the warm ISM simulation, we follow the scheme from Schönrich & Binney (2009a) and inject 99% of stellar yields from all sources (CCSNe, SNeIa, and AGB stars) into the warm ISM with the remainder of the yields entering the cold ISM directly. In each time step, a fraction of the gas \( (dt/t_{cool}) \) cools out of the warm ISM and joins the cold ISM. We adopt a gas cooling timescale, \( t_{cool} \), of 1.2 Gyr. The gas in the warm ISM is not ejected in outflows.

Including a warm ISM delays the return of yields from all sources, in contrast to changing the minimum time delay for SNIa (see Section 4.3.6), which only affects one yield source. In the [O/Fe]–[Fe/H] plane, the model with a warm ISM experiences a slower enrichment and reaches a lower [Fe/H] before turning down at the knee relative to the fiducial simulation. However, the warm ISM simulation is more oxygen-rich than the fiducial simulation above the knee because the warm ISM shields CCSN products like oxygen from ejection and delays iron enrichment from SNeIa.

4.4. Yields and Multi-element Abundances

4.4.1. Yields

The adopted nucleosynthesis yields remain one of the largest sources of uncertainty in chemical evolution models (Romano et al. 2010). The yields of \( \alpha \)-elements from CCSNe and iron-peak elements from SNeIa are relatively robust, but the yields
of explosive and neutron-capture elements and elements expected to have a large metallicity dependence are much more uncertain. Here we discuss our choice of yields and how they affect the chemical evolution model.

**CCSNe**

CCSNe produce the majority of elements heavier than helium, including $\alpha$-elements (O, Mg, Si, S, Ca, and Ti), but there is not yet a consensus on the yield of each element as a function of stellar mass and metallicity. The yields of $\alpha$-elements produced in hydrostatic burning phases (O and Mg) increase significantly with stellar mass, but the yields of $\alpha$-elements produced in explosive burning phases (Si, S, Ca, and Ti) are much less sensitive to stellar mass. The yields of all $\alpha$-elements have a very weak metallicity dependence. On the other hand, the yields of odd-Z elements (odd number of protons), such as sodium and aluminum, increase sharply with metallicity because the initial abundance of carbon and oxygen dictates the neutron fraction in the star. Stars with higher neutron fractions produce larger quantities of odd-Z elements.


---

1Both are available on the ORFEO database at http://www.iasf-roma.inaf.it/orfeo/public_html/.
spanning 8–100 $M_\odot$ at the metallicities of the models by adopting the yields of the closest mass model at fixed metallicity for masses that are higher or lower than those covered by these models (but between 8–100 $M_\odot$). Then we linearly interpolated the yields to create a uniform grid of yields finely sampled in mass and metallicity. Zero metallicity CCSN yields (especially for nitrogen) are highly uncertain, because of the poorly constrained amount of rotational mixing that occurs in zero metallicity stars, so we adopt the lowest metallicity Chieffi & Limongi (2004) yields ($Z = 10^{-6}$) for stars with yet lower metallicity. For super-solar metallicities, we used the solar metallicity Limongi & Chieffi (2006) yields. We linearly extrapolated the remnant mass to higher and lower masses at fixed metallicity.

Chieffi & Limongi (2004) and Limongi & Chieffi (2006) report yields as a function of the mass cut in CCSNe, which is the dividing line between material that falls back onto the neutron star and material that is ejected in the explosion. Since the mass cut is deep in the star, its location affects the yields of Fe-peak elements but not those of elements created in the outer layers of the star like oxygen. As shown in Figure 4.5, a higher mass cut decreases the [O/Fe] of the plateau, increases [Fe/H], and increases the [Fe/H] of the knee. This leads to a significant degeneracy between the mass cut, the SFE (Figure 4.4a), and the SNIa DTD (Figure 4.9). The appropriate mass cut is uncertain at the factor of 2–3 level and might itself change as a function of stellar mass (Limongi & Chieffi 2003). We adopt a mass cut such that 0.1 $M_\odot$ $^{56}$Ni is produced in all CCSNe. This mass cut reproduces
the [O/Fe] abundances for stars with [Fe/H] < −1 (see Figure 4.5b) and is the same mass cut adopted by Chieffi & Limongi (2004) and Limongi & Chieffi (2006). (However, Schönrich & Binney 2009a chose a mass cut of 0.05 M⊙ 56Ni to match the [Ca/Fe] abundances of Lai et al. 2008.) While the [O/Fe] plateau seems like a powerful criterion for selecting the mass cut, there is observational uncertainty in the measured value at the level of ∼0.2 dex. For example, the [O/Fe] plateau at low [Fe/H] shifted from ∼0.6 in Ramírez et al. (2007) to ∼0.45 in Ramírez et al. (2013) using much of the same data due to the adoption of the improved calibration of the temperature scale from Casagrande et al. (2010).

The yields for neutron-capture elements are not well-predicted by theoretical yield calculations. We adopt the “empirical” r-process yields for Eu and Ba from Cescutti et al. (2006), which were determined by fitting a chemical evolution model to Eu and Ba abundances with different yields. The Cescutti et al. (2006) yields do not depend on metallicity. We report the evolution of Eu abundances even though the application of “empirical” yields is partially circular (we use a different chemical evolution model). We do not show the evolution of Ba abundances because of the lack of reliable s-process Ba yields. The s-process contributes significantly to Ba abundances but not Eu abundances (Sneden et al. 2008).
SNeIa synthesize at least half of the iron and iron-peak elements in the Galaxy. They are rarer than CCSNe by a factor of $\sim 4 - 5$ (time-integrated; Maoz et al. 2011) or $\sim 3$ (current rate in local surveys; Mannucci et al. 2005; Li et al. 2011), but they make $\sim 0.6 - 0.7$ M$_\odot$ of $^{56}$Fe whereas CCSNe produce $\sim 0.05 - 0.1$ M$_\odot$ of $^{56}$Fe. Depending on the adopted yields, SNeIa also can create $\alpha$-elements like oxygen, calcium, and titanium, but in quantities that are insignificant relative to the CCSN contribution to the galactic budget of these elements (see Figures 4.2 and 4.12).

We adopted the SNIa yields of the W70 model from Iwamoto et al. (1999). Unlike CCSN yields, chemical evolution is relatively insensitive to the specific choice of SNIa yield model because most models produce similar amounts of iron and iron-peak elements. This is also borne out by observations of SNIa lightcurves, which are powered by the radioactive decay of $^{56}$Ni (eventually into $^{56}$Fe) and constrain the amount of iron produced to be $\sim 0.6$ M$_\odot$ on average. The SNIa models do differ in the quantities of other elements synthesized, but SNeIa are not the dominant contributors of these elements. We assume that SNIa yields are independent of the mass and metallicity of the progenitor stars, which is probably a good assumption given the arguments above.
AGB Stars

AGB stars dominate the production of nitrogen and s-process elements; are important producers of carbon; and return large amounts of hydrogen and helium to the Galaxy (see Figure 4.12). During hydrogen-burning, they convert carbon and oxygen into nitrogen, which is the bottleneck of the CNO cycle. They also synthesize carbon and oxygen, but most of it remains in the core that forms the white dwarf, and the net yield for these elements can even be negative at certain stellar masses and metallicities. AGB stars generate s-process elements in the thermal pulses at the end of their lives, but modeling these pulsations is complex, so predicting s-process yields is challenging. Thus, we do not try to model s-process elements, though they are important chemical tracers of enrichment from low mass stars.

We used the AGB yields from Karakas (2010) that span 1–6.5 \( M_\odot \) and \( 10^{-4} \text{–} 1.0 \) \( Z_\odot \). We linearly interpolated between the grid points in mass and log metallicity, and linearly extrapolated the yields up to 8 \( M_\odot \). CCSNe consume hydrogen but have positive or zero yields for other elements. However, AGB stars can either destroy or produce carbon and oxygen, depending on metallicity and mass, so interpolating and extrapolating the yields of these elements can be uncertain.

4.4.2. Multi-element Abundances

In Figure 4.11 we compare the abundances from two simulations with different CCSN yields, one with Chieffi & Limongi (2004) and Limongi & Chieffi (2006)
yields (the same as the fiducial model; hereafter CL04) and one with the Woosley & Weaver (1995, hereafter WW95) yields. We reduced the WW95 yields of the iron-peak elements (Cr, Mn, Fe, Co, Ni, Cu, and Zn) These simulations have a constant SFR but otherwise adopt the parameters of the fiducial model, including the SNIa and AGB yields described above. Relative to the fiducial simulation, the constant SFR simulation has a lower equilibrium abundance and higher equilibrium $[X/Fe]$ values (except manganese and nickel). Below we compare the predicted abundances of twenty elements for the two yield sets and discuss the agreement or disagreement with data from Reddy et al. (2003, 2006) and Ramírez et al. (2013) (oxygen only). Figure 4.12 shows the mass enrichment histories for these elements in the same format as Figure 4.2. We discuss results element-by-element with reference to both figures.

In general, the two sets of yields produce roughly comparable trends for $\alpha$-elements, whose yields are metallicity-independent, but very different trends for elements with metallicity-dependent yields, like odd-Z elements. The typical $\alpha$-element trend is characterized by a plateau at super-solar abundance at low metallicities ($[Fe/H] \lesssim -1$), where the trend turns over and declines to solar abundance at solar metallicity. The normalization of the plateau is set by the CCSN yields, and the turnover is caused by the onset of SNIa iron enrichment. AGB stars do not significantly affect $\alpha$-element abundances. The CCSN yields of odd-Z elements increase with metallicity, so their predicted abundance trends rise
with increasing metallicity and may turn over due to iron enrichment from SNeIa. Elements with metallicity-dependent yields have a steeper blue solid curve in Figure 4.12 than the red solid curve (iron mass). The WW95 iron yields have a metallicity dependence at low metallicity, which causes the decline in [O/Fe] for [Fe/H]< −1. This feature is present for all elements including [Eu/Fe], whose yield is completely independent of metallicity.

**Carbon**  Carbon enrichment, like oxygen, is dominated by CCSNe. With our adopted yields, the contribution of AGB is sub-dominant but non-negligible above [Fe/H] ≈ −0.6. The [C/Fe] data follow an α-element trend except that [C/Fe] turns over at [Fe/H] = −0.4, whereas the α-elements turn over at [Fe/H] = −0.9. [C/Fe] actually increases at [Fe/H] = −0.7, where carbon production by AGB stars peaks. At higher [Fe/H], AGB stars become net consumers of carbon. The CL04 simulation matches the observational data well for [Fe/H]> −0.4, but it predicts too low an abundance for the plateau. The WW95 simulation produces a similarly shaped trend to the data that is systematically too low by ~0.3 dex.

**Nitrogen**  Most nitrogen production occurs in the CNO cycle, where carbon and oxygen get converted into nitrogen. Figure 4.12 shows that AGB stars dominate the production of nitrogen, especially for [Fe/H]< −1, so it is not surprising that the CL04 and WW95 simulations result in similarly shaped tracks in [N/Fe]–[Fe/H]. The
[N/Fe] of the CL04 and WW95 simulations increases quickly from low metallicity to [Fe/H] \sim -0.6 and -0.8, respectively, which reflects the secondary nature of nitrogen (i.e., its yield increases with the initial carbon and oxygen abundances). This metallicity dependent yield is also evident in the steepness of the blue dashed curve in Figure 4.12. It then declines as the nitrogen production from AGB stars saturates (with zero net yield) while iron production from SNeIa continues to increase (Figure 4.12). Both simulations underpredict [N/Fe] by \sim 0.5 dex over the limited range of the data \(-0.4 \lesssim [\text{Fe/H}] \lesssim 0.1\).

Massive stars, (i.e., ones that will explode as CCSNe) could be an important source of nitrogen at zero and low metallicities. However, the uncertainty in the nitrogen yield of zero metallicity massive stars is believed to at least an order of magnitude (Heger & Woosley 2010). These stars are born with low abundances of carbon and oxygen (relative to hydrogen), but rotational mixing could transport freshly synthesized carbon and oxygen from helium-burning layers to hydrogen-burning layers, making nitrogen production extremely sensitive to the strength of mixing. Nitrogen abundances are also difficult to measure, and super-solar [N/Fe] for stars with only slightly sub-solar metallicity would imply a dramatically different enrichment history for these stars compared to the Sun if these measured abundances are correct. Nonetheless, if observational estimates are approximately correct, then either the AGB or CCSN yields (or both) must be substantially in error.
Oxygen Oxygen is produced almost exclusively by CCSNe (see Figure 4.12). Its yield has little dependence on metallicity but a strong dependence on stellar mass, as can be seen from the effects of changing the slope or cutoff of the high mass end of the IMF in Figure 4.5b. The parameters of the fiducial simulation were chosen to match the data in \([\text{O/Fe}] - [\text{Fe/H}]\) and reach equilibrium at solar oxygen abundance and metallicity, so the CL04 simulation fits the data by construction. It was not intended to fit the Ramírez et al. (2013) data above solar metallicity, as these stars were likely born interior to the solar annulus in the metal-rich inner disk and migrated to the solar neighborhood. The WW95 simulation fits the data well for \([\text{Fe/H}] > -0.4\), though it overpredicts \([\text{O/Fe}]\) at lower metallicities. The differences between the WW95 and CL04 simulations at low metallicity are partially due to the metallicity dependence of the WW95 iron yields, whereas the CL04 iron yields are metallicity-independent by design since we chose a fixed mass cut.

The range of plausible mass cuts and iron yields for CCSNe spans the range of uncertainty in the oxygen abundance measurements, which are challenging, even for the Sun. In particular, abundances determined from the strong OI triplet at 777 nm are sensitive to the adopted temperature scale and require corrections from the typical assumption of local thermodynamic equilibrium, adjustments of 0.1–0.5 dex depending on gravity and temperature (Ramírez et al. 2013).
Sodium  Nearly all sodium is synthesized in CCSNe. Its CCSN yield is strongly metallicity-dependent because it is an odd-Z element, hence the steepness of the blue dotted curve in Figure 4.12 and the rising trend of [Na/Fe] up to [Fe/H]< −0.4 for the CL04 simulation. At higher metallicities, [Na/Fe] decreases due to increased iron enrichment from SNeIa. The CL04 simulation reproduces the observed [Na/Fe] trend, which is flat at [Na/Fe] ≈ 0.1 for [Fe/H]< −0.4 with a slow decline toward solar [Na/Fe] at solar [Fe/H]. The WW95 simulation initially declines from [Fe/H] = −2 to −1.6 due to its metallicity-dependent iron yields. It then rises from [Fe/H] = −1.6 to −0.6 but more gradually than the CL04 simulation due to a weaker metallicity dependence for the sodium yields. At [Fe/H] = −0.6, it turns over because of SNIa iron enrichment. The normalization of the WW95 trend is a better match to the data than the CL04 simulation and does a reasonable job of fitting the data albeit with a sharper turn over. There are some observational uncertainties in measuring sodium abundances, including non-local thermodynamic equilibrium effects.

Magnesium  Magnesium enrichment is dominated by CCSNe, and the predicted yield is only mildly metallicity-dependent. The [Mg/Fe] data follow the typical α-element trend and are well reproduced by the CL04 simulation, though it slightly underpredicts [Mg/Fe], by ∼0.05–0.1 dex. The WW95 simulation also underpredicts [Mg/Fe], by ∼0.15–0.2 dex. François et al. (2004) also found that magnesium is
underproduced by the WW95 yields, and they noted that magnesium yields depend on the treatment of convection.

**Aluminum** Aluminum is an odd-$Z$ element produced almost entirely by CCSNe. The CL04 and WW95 simulations produce aluminum trends that are similar to those of sodium because the physics of production is similar. The CL04 simulation does a poor job of fitting the observational data for $[\text{Al}/\text{Fe}]$, which resembles the $\alpha$-element trend. The CL04 simulation overpredicts $[\text{Al}/\text{Fe}]$ above $[\text{Fe}/\text{H}] = -0.4$ by $\sim 0.15$ dex but dramatically underpredicts it at low metallicity. The weaker metallicity dependence of the WW95 yields relative to the CL04 yields results in a flatter $[\text{Al}/\text{Fe}]$ trend that provides a better fit to the data. It underpredicts the normalization of the observations by $\sim 0.1$ dex, but its shape is consistent with the data for $[\text{Fe}/\text{H}] > -1$.

**Silicon** Silicon is an $\alpha$-element made mostly in CCSNe like oxygen and magnesium but with a larger, though still small, contribution from SNeIa. Both the CL04 and the WW95 simulations do an excellent job of reproducing the observed $[\text{Si}/\text{Fe}]$ data. François et al. (2004) found good agreement with data when they adopted the solar metallicity WW95 silicon yields.

**Sulfur** Sulfur is also an $\alpha$-element, and its production is very similar to that of silicon with a slightly smaller contribution from CCSNe. The CL04 and WW95
simulations agree with the [S/Fe] measurements, though they span almost 0.5 dex in [S/Fe] and only 0.8 dex in [Fe/H].

Potassium  Potassium is an odd-Z element produced almost exclusively in CCSNe. Its yield is predicted to be metallicity-dependent, as can be seen from the super-linear slope of the blue dotted line in Figure 4.12 though its metallicity dependence is weaker than sodium or aluminum. The CL04 simulation significantly underpredicts the amount of potassium needed to achieve the solar value of [K/Fe] (the blue solid line is always well below the red solid line in Figure 4.12) and the observed [K/Fe] data in Figure 4.11 (the CL04 curve falls off the bottom of the plot but the highest value it reaches is [K/Fe] = −0.65). The WW95 simulation also falls far below the data, though by 0.6 dex. The super-solar [K/Fe] values suggest that potassium may behave like an α-element at lower metallicities but more data are needed to confirm such a trend.

Calcium  Calcium is an α-element made mainly in CCSNe, but with a significant SNeIa contribution. The CL04 simulation matches the level of the plateau in the [Ca/Fe] data, but its track turns over more sharply than the data and ends at a sub-solar [Ca/Fe] value. The WW95 simulation produces higher [Ca/Fe] values at low metallicity but it overlaps with the CL04 track for [Fe/H] > −0.8.
**Scandium**  CCSNe are the dominant source of scandium. Like other odd-$Z$ elements, the CL04 scandium yields have a very strong metallicity dependence that results in a rising trend with $[\text{Fe/H}]$, which turns over at $[\text{Fe/H}] \approx -0.3$. However, the observed scandium abundances follow the typical $\alpha$-element trend. The WW95 simulation predicts an $\alpha$-element trend for scandium, in stark contrast to the CL04 simulation. Both the CL04 and WW95 simulations underpredict the overall normalization of $[\text{Sc/Fe}]$. The offset for the WW95 simulation is $\sim 0.3$–0.4 dex. A generic offset cannot be determined for the CL04 simulation because of its dramatically different shape compared to the data, though its peak is below the main trend by $\sim 0.15$ dex.

**Titanium**  Titanium is both an $\alpha$-element and a low iron-peak element. It is mostly produced in CCSNe, though SNeIa contribute about 1/3 of the solar titanium abundance. Both the CL04 and the WW95 simulations produce $\alpha$-element trends with similar normalizations. The data also follow an $\alpha$-element trend. However, both simulations underpredict the trend of the data by $\sim 0.3$–0.4 dex at all metallicities. The underproduction of titanium is a generic problem for SN yields, and its cause is not well understood.

**Vanadium**  Vanadium is an odd-$Z$ element produced predominantly in CCSNe but with a significant SNeIa contribution. The CL04 vanadium yields have a weak
metallicity dependence, and the CL04 simulation predicts a slightly rising [V/Fe] trend with [Fe/H] that peaks at [V/Fe] \( \approx -0.4 \). However, the [V/Fe] data show a weak \( \alpha \)-element trend that declines from [V/Fe] \( \sim 0.2 \) to \( -0.1 \). The WW95 simulation produces an \( \alpha \)-element trend like the data, but it also underpredicts the normalization of the trend by \( \sim 0.4\)–0.5 dex.

**Chromium** Chromium is an iron-peak element made equally in CCSNe and SNeIa. The CL04 yields produce slightly super-solar [Cr/Fe], which declines to slightly sub-solar [Cr/Fe] due to the sub-solar Cr/Fe SNeIa yield ratio. The measured [Cr/Fe] data are centered on solar [Cr/Fe] with little scatter down to [Fe/H] = \(-1\). The WW95 simulation declines by \( \sim 0.2 \) dex from [Fe/H] = \(-2 \) to \(-1\) and then rises slightly. It underpredicts the data by about 0.1 dex.

**Manganese** SNeIa and CCSNe contribute to the enrichment of manganese, an iron-peak element, at almost the same level. The observed data show a unique trend of a plateau at [Mn/Fe] \( \approx -0.4 \) for [Fe/H] < \(-1\) that rises to solar [Mn/Fe] at solar metallicity. the CL04 simulation reproduces the data nearly perfectly, and the WW95 simulation reproduces the shape of the observed [Mn/Fe] trend with a minor offset to lower [Mn/Fe]. The manganese yields are sensitive to the adopted mass cut, in a way that is not offset by similar changes to the iron yields. However, Bergemann & Gehren (2008) found that the observed trend is caused by erroneously
assuming local thermodynamic equilibrium and the trend goes away if non-local thermodynamic equilibrium effects are taken into account.

**Cobalt**  
Cobalt is an iron-peak element whose CCSNe contribution is three times its SNeIa contribution. The CL04 cobalt yields are metallicity-dependent, which produces the rising [Co/Fe] trend with increasing [Fe/H]. In contrast to the CL04 simulation, the observed [Co/Fe] data follow an $\alpha$-element trend. This discrepancy is also present when comparing the data to the WW95 simulation. The WW95 simulation predicts a nearly identical trend to the CL04 simulation for $[\text{Fe/H}] < -0.8$, but it turns over at $[\text{Fe/H}] = -0.6$, whereas the CL04 simulation continues to rise until $[\text{Fe/H}] = -0.3$.

**Nickel**  
Nickel is an iron-peak element that has slightly more contribution from SNeIa than CCSNe. The CL04 and SNeIa nickel yields are metallicity-dependent, so the CL04 simulation predicts a monotonically rising [Ni/Fe] trend. However, the observations show a flat trend at solar [Ni/Fe] with little scatter. The WW95 simulation also fails to reproduce the data. It predicts a similarly shaped trend to the CL04 simulation that is offset towards higher [Ni/Fe] by $\sim 0.1$–0.2 dex.

**Copper**  
Copper is an iron-peak element that is exclusively produced in CCSNe. Both the CL04 and the WW95 copper yields have a strong metallicity dependence, and they predict rising [Cu/Fe] trends that turn over when SNeIa begin to contribute
significant amounts of iron. The [Cu/Fe] measurements show a similar trend to the simulations, though the mean trend is offset to higher [Cu/Fe] by \( \sim 0.1 \) dex. The data have significant scatter, especially at low metallicity. The WW95 simulation fits the low metallicity data better than the CL04 simulation because the former is offset towards lower [Fe/H].

**Zinc**  Zinc is an iron-peak element, whose yields are dominated by CCSNe. Both simulations produce an undulating trend that first declines, then rises due to a weak metallicity dependence of the zinc yields, and finally declines because of SNeIa iron enrichment. The observed [Zn/Fe] data show a weak \( \alpha \)-element trend with large scatter below [Fe/H] = \(-1\). The shape of the simulated trends may be consistent with the shape of the observed trend. However, the overall normalization of the WW95 and CL04 trends are too low by \( \sim 0.25 \) dex and \( \sim 0.5 \) dex, respectively, indicating that the yields are grossly underproducing zinc.

**Europium**  Europium is produced almost exclusively in the \( r \)-process. The astrophysical site of the \( r \)-process is unknown, but the two most likely candidates, CCSNe and neutron star mergers, are associated with massive stars. The [Eu/Fe] data follows an \( \alpha \)-element trend, which strongly hints that europium yields are independent of metallicity. Neither the CL04 nor the WW95 yields included europium, so we adopted the Cescutti et al. (2006) Eu yields (their model 1),
which are metallicity-independent by construction. The differences between the two simulations result from differences in the CCSNe iron yields. The Cescutti et al. (2006) europium yields were determined empirically by fitting a chemical evolution model to observed [Eu/Fe] data from [Fe/H] = −3.4 to +0.15. Given the empirical nature of the europium yields, it is not surprising that the simulations and the data agree, but it is reassuring that we fit a different observational [Eu/Fe] data set than the calibrating one with a different chemical evolution model.

4.5. Bimodality in [α/Fe]–[Fe/H]

The distribution of stars in [α/Fe]–[Fe/H] is bimodal with a high-α population and a low-α population (e.g., Bensby et al. 2003, 2005, 2014; Reddy et al. 2003, 2006; Fuhrmann 2011). The high-α population runs along the low metallicity plateau at super-solar [α/Fe], turns over at [Fe/H] ∼ −0.8, and reaches solar [α/Fe] at super-solar metallicity. The low-α population starts at solar or slightly super-solar [α/Fe] ([O/Fe] ∼ 0.2 for the Ramírez et al. 2013 data) and [Fe/H] ∼ −0.5 and runs through solar [α/Fe] and [Fe/H]. The low-α population has a shallower slope than the high-α population for [Fe/H] > −0.5, causing them to merge at solar [α/Fe] and super-solar [Fe/H].

The observed frequency of stars with intermediate [α/Fe] is uncertain because many surveys (e.g., Bensby et al. 2003, 2005; Reddy et al. 2003, 2006) use kinematic selection criteria to target thick disk stars that likely have high [α/Fe] and avoid
stars with intermediate kinematics that may be scattered thin disk stars. Only selecting stars with extreme kinematics produces the cleanest separation between the high- and low-\(\alpha\) populations, but it exaggerates the underdensity at intermediate \([\alpha/Fe]\). In an effort to find stars with intermediate \([\alpha/Fe]\), Ramírez et al. (2013) and Bensby et al. (2014) targeted stars with intermediate kinematics and found that stars with intermediate kinematics can have high-, intermediate-, or low-\(\alpha\) abundances, so kinematically selecting intermediate \([\alpha/Fe]\) stars remains challenging.

Issues with target selection are partially mitigated by volume-limited surveys, such as Fuhrmann (1998, 2004, 2008, 2011). This sample of 271 stars shows a gap in \([\text{Mg}/\text{Fe}] - [\text{Fe/H}]\), but it only contains 15 thick disk (high \([\text{Mg}/\text{Fe}]\)) and 5 intermediate \([\text{Mg}/\text{Fe}]\) stars because the local nature \((d < 25 \text{ pc})\) of this sample limits its ability to probe the thick disk. Adibekyan et al. (2012) analyzed a larger sample of 1111 stars targeted by HARPS that was not strictly volume-limited but was chosen without kinematic cuts. The valley at intermediate \([\alpha/Fe]\) for the HARPS sample becomes less populated when only the stars with the most accurate abundance measurements are considered (stars with \(T_{\text{eff}} = T_{\text{eff, \odot}} \pm 300 \text{ K}\)), suggesting that observational errors may be scattering stars into the valley. Using SEGUE data, Bovy et al. (2012) found that the distribution of observed stars is clearly bimodal, but once they account for sample selection effects the distribution has a low-\(\alpha\) peak with a tail to high-\(\alpha\).

Initial results from APOGEE (Anders et al. 2014; Hayden et al. 2014) find that the
gap is depopulated but not completely empty and becomes more underdense with
decreasing metallicity.

Two scenarios for creating the bimodality are the two infall model of Chiappini et al. (1997, 2001) and the Schönrich & Binney (2009a,b) model that relies on
mixing stellar populations born at different radii. While these two models can both reproduce the bimodality, they make different predictions for the density of stars in the valley. The two infall model predicts a true gap because of the gap in the SFH, whereas the Schönrich & Binney (2009a,b) predicts a small but non-zero number of stars in the valley.

4.5.1. Two Infall Model

The low-\(\alpha\) population typically has been thought of as an evolutionary sequence.
However, connecting the high-\(\alpha\) population to the low-\(\alpha\) population is challenging for chemical evolution models. Chiappini et al. (1997, 2001) found that a track could run through both populations if the disk formed in two major infall episodes
separated by a cessation of star formation caused by a threshold gas density for star formation (Kennicutt 1989). Figure 4.13a shows a pair of simulations of the two infall model for hiatuses in star formation of 200 Myr (blue points) and 1 Gyr (red points). For the first Gyr, we set the inflow timescale to be 1 Gyr and set the SFE to be \(6 \times 10^{-10}\) Gyr\(^{-1}\). Next, we completely suppressed star formation starting at 
\(t = 1.0\) Gyr for 200 Myr or 1 Gyr. Then, we set the inflow timescale to be 6 Gyr and
reduced the SFE by a factor of two (to $3 \times 10^{-10}$ Gyr$^{-1}$) with a maximum inflow ($\tau_{\text{max}}$) at 1 Gyr ($\propto e^{-\left(t-\tau_{\text{max}}\right)/\tau_1}$).

The two infall model produces a true gap in [O/Fe]–[Fe/H] because of the hiatus in star formation. During the hiatus, SNIa iron enrichment decreases [O/Fe]–[Fe/H] but its effect of increasing [Fe/H] is offset by dilution from inflow, resulting in only a slight net decrease in [Fe/H]. The magnitude of the gap depends sensitively on the length of the hiatus in star formation: 0.07 vs. 0.23 dex in [O/Fe] for the 200 Myr and 1 Gyr hiatuses, respectively. Since the large gap and inverted-U shape of the low-α population of the 1 Gyr hiatus simulation are not seen in the data, the length of the hiatus must fall in a narrow range around 200 Myr to produce the right size gap and the morphology of the low-α population.

4.5.2. Mixing Stellar Populations

The bimodality in [O/Fe]–[Fe/H] could also be the result of dynamical processes mixing stars from different birth radii and hence with different enrichment histories. In the case of local super-solar metallicity thin disk stars, Grenon (1987) and Francois & Matteucci (1993) argued that these stars were born in the inner disk and migrated to the solar neighborhood by diffusion. More recently, Sellwood & Binney (2002) discovered another mixing process called radial migration that resonantly scatters stars off of spiral arms and transports them large distances without inducing a large ellipticity, so their kinematics resemble locally born thin disk stars.
Schönrich & Binney (2009a,b) incorporated diffusion and radial migration into a chemical evolution model that naturally produced the high-α and low-α populations from a superposition of stars born at a range of radii without a hiatus in star formation. In their model, the high-α population is formed by overlapping tracks from different radii. Once SNeIa become significant contributors of iron, the tracks move quickly across the valley in $[\alpha/Fe]–[Fe/H]$. Their model produces some but not many intermediate $[\alpha/Fe]$ stars—in contrast to the two infall model, which produces none. They find that the low-α population is a superposition of the stars near the equilibrium abundances of a range of radii and not an evolutionary sequence. Since individual tracks rapidly asymptote to their equilibrium abundance, the low-α population contains many more stars than the valley.

Figure 4.13b shows a superposition of simulations that was designed to reproduce the bimodality in $[O/Fe]–[Fe/H]$ and the low-α population by varying the outflow mass-loading parameter and inflow timescale to mimic the enrichment histories of several different galactocentric radii, whose stars were then mixed together. Increasing $\eta$ decreases the equilibrium $[Fe/H]$ (see Figure 4.4b) and results in a range of equilibrium metallicities that form a ridge line at low-α. Increasing the inflow timescale maintains the same trajectory of a simulation but moves its equilibrium abundance higher up the track to higher $[O/Fe]$ and lower $[Fe/H]$, which produces the negative slope of the low-α population. To reproduce the slope, we adopted a constant SFR for the highest-$\eta$ simulation that effectively acts as an
infinite inflow timescale. As the inflow timescale and $\eta$ increase, the final stellar mass of the simulations decrease (by a factor of three from the lowest-$\eta$ to the highest-$\eta$ simulations) because the simulations accrete less gas and retain that gas less efficiently. The colored points are stars randomly drawn from the simulations with Gaussian noise of $\sigma = 0.05$ in [Fe/H] and $\sigma = 0.02$ in [$\alpha$/Fe] added to display stars with identical abundances and for ease of comparison with the data.

In Figure 4.13b, the number of stars shown from each simulation is proportional to its stellar mass, but the effectiveness of stellar mixing decreases with distance, which would decrease the number of stars from the most distant and hence most metal-rich and metal-poor simulations. This suite of simulations is intended to illustrate how the superposition scenario works, and a more complete multi-zonal model with an accurate treatment of migration is needed to make more quantitative comparisons with data.

While these parameter variations were motivated by the results shown in Figures 4.3a and 4.4b, they reflect differences in formation history as a function of galactocentric radius. An increasing inflow timescale with radius is a generic aspect of inside-out galaxy formation. The outflow mass-loading parameter also should increase with radius because gas and metals flow inwards through the disk (Stark 1984), enriching the inner disk at the expense of the outer disk. In addition, the outer disk may be less effective at retaining gas and metals due to a weaker vertical potential and a lower gas density to counteract energy injection by SNe.
To summarize, a single simple chemical evolution model cannot reproduce the bimodal $\alpha$-element distribution. The two infall model creates a bimodal distribution and produces a gap between the high-$\alpha$ and low-$\alpha$ populations due to a gap in the SFH. The length of this hiatus in star formation must be narrowly confined to about 200 Myr to avoid overshooting the low-$\alpha$-population. Alternatively, the bimodality can be produced by a mix of stellar populations born at different galactocentric radii with unique enrichment histories that migrated to the solar circle. This superposition scenario partially populates the valley in [$\alpha$/Fe], meaning that large future data sets with better understood selection effects may be able to differentiate between these two scenarios based on the intrinsic level of depopulation in the valley.

4.6. Conclusions

Chemical evolution models are a powerful tool for interpreting stellar abundances in the context of galaxy evolution. However, the results of such models can depend heavily on the treatment of galaxy evolution and enrichment processes. We constructed a one-zone open box chemical evolution model to investigate the effect of variations in the inflow timescale, the functional form of the inflow rate, the SFE, the outflow mass-loading parameter, the inflow metallicity, the IMF, the CCSN mass cut, the minimum SNIa delay time, the SNIa delay time distribution, the inclusion of a warm phase of the ISM, and the CCSN yields. The two main features in [O/Fe]–[Fe/H] are the knee, which is affected by SFE, SNIa DTD, and a warm
ISM, and the equilibrium abundance, which is governed by the outflow mass-loading parameter and the inflow timescale. The IMF and CCSN mass cut can dramatically shift the location of the trajectory in [O/Fe]–[Fe/H] but changes in the other model parameters can compensate for these shifts. The choice of CCSN yields strongly impacts the predicted evolution of elements with metallicity-dependent yields, such as odd-Z elements. To explain the bimodality in [O/Fe]–[Fe/H] we compare the two infall model and a superposition scenario, which is a mix of stellar populations born at different radii. The low-\(\alpha\) population of the latter scenario is not an evolutionary sequence but rather a superposition of the equilibrium abundances from different radii. The superposition scenario produces an underdensity at intermediate-\(\alpha\), in contrast to the two infall model that predicts a true gap in [\(\alpha/\text{Fe}\)], whose magnitude is very sensitive to the length of the gap in the SFH.
Fig. 4.1.— [O/Fe]–[Fe/H] for a closed box (no gas inflow or outflow) simulation, an inflow only simulation, and the fiducial model that includes both inflow and outflow. Inflow dilutes the metallicity but boosts [O/Fe] because it fuels late time star formation and CCSN enrichment. Outflow decreases metallicity by ejecting metals and the gas that would have fueled additional star formation. Outflow also suppresses [O/Fe] because it has more opportunities to remove elements produced earlier, like O, than those made later, like Fe. The black solid lines mark the solar abundance. Time is indicated by the colored circles (100 Myr), squares (500 Myr), upward-pointing triangles (1 Gyr), diamonds (2 Gyr), and downward-pointing triangles (4 Gyr). The colored tick marks on the right indicate the equilibrium [O/Fe]. The gray points show the iron abundances and NLTE oxygen abundances of 775 local thin disk, thick disk, and halo stars from Ramírez et al. (2013).
Fig. 4.2.— The mass of gas-phase oxygen and iron as a function of [Fe/H] for the constant SFR simulation. The line styles indicate the contributions from different enrichment sources (including recycled material—i.e., atoms incorporated into a star at birth and returned at death). The ordinate axes are aligned such that the total oxygen and total iron lines cross at solar [O/Fe]. The black bar indicates a 0.5 dex offset. We show the constant SFR simulation because its iron mass is linear with [Fe/H], facilitating a comparison between oxygen and iron.
Fig. 4.3.— $[\text{O}/\text{Fe}]$–$[\text{Fe/H}]$ for variations in the inflow timescale (panel a) and the functional form of the inflow rate (panel b). Symbols and data points are the same as in Figure 4.1.

Fig. 4.4.— $[\text{O}/\text{Fe}]$–$[\text{Fe/H}]$ for variations in the star formation efficiency and the outflow mass-loading parameter $\eta$ (see Equation 4.5). Increasing the star formation efficiency increases the $[\text{Fe/H}]$ of the knee. Increasing the outflow mass-loading parameter steepens the trajectory of the trend in $[\text{O}/\text{Fe}]$–$[\text{Fe/H}]$. Symbols and data points are the same as in Figure 4.1.
Fig. 4.5.— [O/Fe]–[Fe/H] for variations in the inflow metallicity and the stellar initial mass function (and the mass cut for CCSNe). We note that panel (b) has a different scale. Symbols and data points are the same as in Figure 4.1.
Fig. 4.6.— SFHs for exponential, double exponential, and linear-exponential product inflow rates and constant SFR simulation.
Fig. 4.7.— [O/Fe]–[Fe/H] for the minimum SNIa delay time. Increasing the minimum SNIa delay time increases the [Fe/H] of the knee. Symbols and data points are the same as in Figure 4.1.
Fig. 4.8.— Panel (a): the exponential, power law, and single degenerate scenario SNIa delay time distributions (DTDs; in units of $N_{\text{SNIa}} \, \text{yr}^{-1} \, \text{M}_\odot^{-1}$) for a stellar population from Schönrich & Binney (2009a), Maoz & Mannucci (2012), and Greggio (2005), respectively. The DTDs were normalized to $2.2 \times 10^{-3}$ SNIa/$\text{M}_\odot$ integrated from 40 Myr to 10 Gyr. Panel (b): number of SNeIa as a function of time (for the same SFH) for the various SNIa DTDs, including the prompt + delayed model of Scannapieco & Bildsten (2005).
Fig. 4.9.— [O/Fe]–[Fe/H] for various SNIa DTDs (see Figure 4.8). The exponential, power law, and single degenerate scenario DTD simulations are broadly consistent with the data. The power law DTD simulation extends to higher metallicity and lower [O/Fe] than the fiducial simulation, and the opposite is true for the single degenerate scenario DTD simulation. The prompt + delayed DTD simulation does not fit the data because it produces too much iron at early times, driving down [O/Fe]. Symbols and data points are the same as in Figure 4.1.
Fig. 4.10.— [O/Fe]–[Fe/H] for a simulation with a warm phase of the ISM that does not form stars. In this simulation, 99% of yields from all sources are injected into the warm ISM, and this material returns to the cold ISM on a timescale of 1.2 Gyr, where it can be reincorporated into stars. The other 1% of the yields go directly into the cold ISM. The warm ISM simulation enriches more slowly, produces much lower metallicity stars, and has a lower [Fe/H] of the knee than the fiducial simulation. Symbols and data points are the same as in Figure 4.1.
Fig. 4.11.— $[X/Fe]$–$[Fe/H]$ for CL04 and WW95 yields with a constant SFR. The gray points show data from Ramírez et al. (2013) (oxygen only) from Reddy et al. (2003, 2006) (all other elements).
Fig. 4.12.— The mass of various elements and iron in the gas-phase as a function of [Fe/H] for the constant SFR simulation. Note that the iron enrichment history (red lines) is the same in all panels. The line styles are the same as Figure 4.2.
Fig. 4.13.— Two scenarios for creating the bimodality in [O/Fe]–[Fe/H]. Panel (a) shows two versions of the two infall model of Chiappini et al. (1997, 2001) with hiatuses in star formation of 200 Myr (blue points) and 1 Gyr (red points). The two infall model predicts a true gap in [O/Fe] due to the gap in the SFH. The size of the gap in [O/Fe] is sensitive to the length of the gap in the SFH: a 200 Myr hiatus matches the magnitude of the gap but a 1 Gyr hiatus overshoots it. Panel (b) shows the scenario where the bimodality is reproduced by a superposition of simulations with variations in the outflow mass-loading parameter ($\eta$) and inflow timescale that correspond to enrichment histories from different radial zones. Although these stars were born at different radii, they are transported to the solar circle by diffusion and radial migration. The superposition scenario predicts an underpopulated but not totally empty valley and a low-$\alpha$ population that is not an evolutionary sequence but rather a ridge line formed from the equilibrium abundances over a range of radii. In both panels, the colored points represent stars randomly selected from the simulations with Gaussian noise of $\sigma = 0.05$ in [Fe/H] and $\sigma = 0.02$ in [O/Fe] added. The open symbols and gray data points are the same as in Figure 4.1.
References

Bensby, T., Yee, J. C., Feltzing, S., et al. 2013, aap, 549, A147

164


De Robertis, M. M., Dufour, R. J., & Hunt, R. W. 1987, JRASC, 81, 195


—. 2004, Astronomische Nachrichten, 325, 3


Garnett, D. R., Edmunds, M. G., Henry, R. B. C., Pagel, B. E. J., & Skillman, E. D.
2004a, AJ, 128, 2772
Grenon, M. 1987, Journal of Astrophysics and Astronomy, 8, 123
Guseva, N. G., Izotov, Y. I., Stasińska, G., Fricke, K. J., Henkel, C., & Papaderos,
Guseva, N. G., Papaderos, P., Izotov, Y. I., Green, R. F., Fricke, K. J., Thuan,
Guseva, N. G., Papaderos, P., Izotov, Y. I., Noeske, K. G., & Fricke, K. J. 2004,
A&A, 421, 519
168


169
Maoz, D., & Mannucci, F. 2012, PASA, 29, 447
O’Connell, D. J. K. (ed.) 1958, Stellar Populations (Vatican Observatory, Vatican City & North Holland Publ., Amsterdam)
Oort, J. H. 1926, Ph.D. Thesis,

173


Peimbert, M., & Costero, R. 1969, Boletin de los Observatorios Tonantzintla y Tacubaya, 5, 3


Searle, L., & Sargent, W. L. W. 1972, apj, 173, 25
Skillman, E. D., Côté, S., & Miller, B. W. 2003, AJ, 125, 610
Talbot, R. J., Jr., & Arnett, W. D. 1971, apj, 170, 409

176


Tinsley, B. M., & Larson, R. B. G., D. Campbell 1977, Evolution of Galaxies and Stellar Populations,


Webster, B. L., & Smith, M. G. 1983, MNRAS, 204, 743