Spectroscopy of Galaxies: Evolution of Escape Fractions, Metallicity Gradients and Stellar Metallicity

Thesis by
Nicha Leethochawalit

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
Defended May 15, 2019

Nicha Leethochawalit
ORCID: 0000-0003-4570-3159

ACKNOWLEDGEMENTS

Six years at Caltech is the longest that I have ever spent at any school. It did not feel long, yet enough to become a scientist with the confidence to do science. Of course, place and time were merely facilitative; it was really the people that defined my time at Caltech.

I was fortunate to have not one but the two best advisors I could have asked for – Evan Kirby and Richard Ellis. Richard adopted me in my early years as a graduate student. Being the luminary to whom I would look for guidance and inspiration, he opened my eyes to the broad spectrum of topics in galaxy evolution. I had the privilege of being checked on by Richard even after he moved to Europe. My heartfelt gratitude also goes to Evan, under whom I spent most of my years in graduate school maturing as a scientist. Besides their performance, Evan also cares deeply about his students' well-being. He gives his students space to explore science independently, but is also always extending his advice and help. I thank Evan for always being available; I was never rebuffed whenever I popped into his office with questions. To both Richard and Evan, I am grateful for your trust in me to do science.

Special thanks to Tucker Jones who taught me how to observe, analyze data and write papers. Also to staff at Caltech, Keck and Palomar (especially Gita, Judy, Carl, Anu, and Patrick) for helping me along.

A shout-out to friends and colleagues! Thanks Yi and Michael E. who always magically had answers to the physics/computer problems that I encountered. Jake, amazing officemate, completing postdoc applications together with you made the process tolerable. Marin, you were like a sister in the department. The lunch cohort, I appreciate the daily department lunch train. My Thai friends, you were a pseudo family. A special mention also to my classmates (Donal and Gina). Lastly, my dear climbing mates, especially Tung, Karen, and Ashish, my life at Caltech would not have been as fulfilling or passable without any of you.

I would like to also extend my gratitude to the people beyond Caltech who have helped me to this point: all my science and physics teachers; mentors at UChicago – Will High and John Carlstrom; and the Thai government for giving me the opportunity to study abroad. Most importantly, I was only able to make it to and through Caltech because of the love that my family and Zhen Hao have generously showered on me.

ABSTRACT

We use spectroscopic observations to investigate galaxies from the age of reionization to the peak of star formation, and the local universe. This thesis presents three projects to better understand characteristics of galactic feedback, i.e., how it regulates galactic gas flows.

We used deep absorption line spectroscopy to estimate the escape fractions $f_{\rm esc}$ of star-forming gravitationally-lensed galaxies at $z \simeq 5$. The approach is to measure the covering fraction of neutral hydrogen in a galaxy from the amount of non-ionizing UV radiation absorbed by low-ionization metal species. With the boost of signal by gravitational lensing, we observed four $4 < z < 5$ star-forming galaxies with DEIMOS, doubling the sample size of existing observations at the redshift range. We found that the escape fractions across our sample varies from galaxy to galaxy and appear to have no significant evolution over time. We inferred a median absolute escape fraction of Lyman continuum photons of $19 \pm 6\%$. Accounting for possible biases and uncertainties, the absolute escape fraction could be reduced to no less than $f_{\rm esc,abs} \sim 11\%$. Moreover, we spatially resolved and detected variation of escape fraction within a galaxy for the first time. The significant variations within the galaxy suggest that the escape fraction is governed by small-scale structure. We found a tentative anti-correlation between the star-formation rate (local or integrated) and the inferred escape fraction. This supports that the escape fraction is associated with the delay time after an episode of star formation. Feedback seems to be more effective in governing both a low SFR and a smaller HI covering fraction.

Spatially resolved spectroscopic observations of galaxies at the peak era of star-formation activities are effective in providing insight into the primitive disks. Specifically, we used metallicity gradients to provide constraints on the amount and extent of feedback produced in star-forming galaxies. We observed 15 star-forming galaxies at $z \sim 2$ with OSIRIS to obtain their kinematic properties and gas-phase metallicity gradients. With helps from AO correction and gravitational lensing, the typical spatial resolution in our study is less than a half-light radius of a typical L^* galaxies at $z \simeq 2$. Combining with the sample in T. A. Jones, R. S. Ellis, et al. (2013), we approximately tripled the existing metallicity gradient measurements. We found a lower fraction of rotationally-supported systems than reported from larger kinematic surveys with coarser spatial resolution, which might be partially due to a our improved spatial resolution. We demonstrated that a high

spatial resolution is crucial for an accurate diagnosis of the kinematic properties and dynamical maturity of $z \simeq 2$ galaxies.

As for metallicity gradients, we found a much higher fraction of $z \simeq 2$ galaxies having weak or flat metallicity gradients than in previous studies. We correlated the metallicity gradient with the total metallicity and found that all galaxies with low total metallicities have flat gradients (< 0.1 dex kpc^{-1}). For galaxies with high metallicities ([N II]/Hα > 0.1), there is a divergence between isolated or rotationally-supported and dynamically-immature systems with the latter showing zero gradients irrespective of the integrated metallicity. The results indicate that relatively strong feedback (e.g. high mass loading factors or high SN energy output) is required in order to explain the majority of the observed flat gradients.

In the second part of the thesis, we observed quiescent galaxies at $z < 1$ and archaeologically constrained galactic feedback via its cumulative influence on stellar metallicities. First, we present the stellar mass–[Fe/H] relationship in the Cl0024+17 galaxy cluster at $z \sim 0.4$. We derived the metalliticies via full spectrum stellar population synthesis modeling of individual quiescent galaxies. Our results provide the metallicities of the lowest galaxy mass ($M_* = 10^{9.7} M_\odot$) at which individual stellar metallicity has been measured beyond the local universe. We found that the mass–[Fe/H] relationship evolves with the redshift at which the galaxy is observed today. Furthermore, we found an even stronger evolution of the mass–[Fe/H] relation with the time at which the galaxy formed (rather than the time at which it is observed). Galaxies that formed earlier have lower Fe abundance than galaxies that formed later.

Lastly, we measured magnesium (Mg) abundances and extended the observed redshift to $z \sim 0.55$. We found that while the mass-[Fe/H] relation evolves significantly over the observed redshift range, the mass–[Mg/H] relation does not. This is due to the shorter star formation histories of quiescent galaxies at higher redshifts. Fe is mainly produced in Type Ia SN. It has a longer recycling time than Mg, which is mainly produced in core-collapse SN. Using core-collapse SN elements as a metal indicator lessens the complication of delayed recycling time and allows us to effectively use galactic chemical models with instantaneous recycling to quantify average outflows that these galaxies experience over their lifetime. We found that the average mass-loading factor η is a power-law function of galaxy stellar mass, $\eta \propto M_*^{-0.21 \pm 0.09}$, consistent with the results of other observational methods and with the predictions where outflow is caused by star formation feedback in turbulent disks.

PUBLISHED CONTENT AND CONTRIBUTIONS

Leethochawalit, N., Kirby, E. N., Ellis, R. S., Moran, S. M., & Treu, T., Submitted to ApJ,
N.L. participated in the conception of the project, collected additional data, reduced and analyzed the data, and wrote the manuscript.

Leethochawalit, N., Kirby, E. N., Moran, S. M., Ellis, R. S., & Treu, T. (2018), ApJ 856, 15, doi: 10.3847/1538-4357/aab26a
N.L. participated in the conception of the project, reduced and analyzed the data, and wrote the manuscript., 15

Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., Richard, J., et al. (2016), ApJ 820, 84, doi: 10.3847/0004-637X/820/2/84
N.L. participated in the observation, reduced and analyzed the data, and participated in the writing the manuscript., 84

Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. (2016), ApJ 831, 152, doi: 10.3847/0004-637X/831/2/152
N.L. participated in the observation, reduced and analyzed the data, and wrote the majority of the manuscript., 152

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Number *Page*

LIST OF TABLES

Chapter 1

INTRODUCTION

Why do galaxies have diverse properties, such as colors, luminosities, dynamics and chemical compositions? Understanding these properties has long been central to the study of galaxy evolution. In the past few decades, we have made great progress in simulating and understanding galaxies in the context of dark matter. Many challenges remain in the poorly understood physics of the baryons, e.g., star formation, turbulence, and outflows. Most of these problems are related to momentum and energy input from stars and massive black holes, i.e., "feedback" processes (for theoretical reviews, see Somerville and Davé, 2015; Naab et al., 2017).

This thesis aims to better understand stellar feedback using direct observations of galaxies throughout the cosmic time. We study three aspects of galaxy evolution: escape fractions of ultraviolet (UV) ionizing photons from high-redshift galaxies, metallicity gradients of star-forming galaxies, and stellar metallicities of quiescent galaxies. First, we provide a brief overview of cosmic history, split into 3 periods: the first billion years, the peak of star-formation, and the quenching period. The details and the main unknowns of each period according to the topics studied in this thesis are in the subsequent sections. Lastly, we list specific questions that this thesis attempts to answer in Section 1.4.

When the universe's recombination epoch ended at $\sim 400,000$ years after the Big Bang, it was mostly dark and composed of neutral hydrogen. Theoretical work suggests that the first stars and dwarf galaxies started to form at $z \sim 30$ (Bromm et al., 2011). Within the first billion years, the neutral hydrogen in the intergalactic medium (IGM) became ionized. Currently, the best constraints on the time and duration of reionization are mainly from studies of high-redshift quarsars and Lyman alpha (Lyα) emitting galaxies (e.g., F. B. Davies et al., 2018; Mason, Treu, et al., 2018). They indicate that most of the reionization occurs at $z \sim 6-8$. Despite the timeline, one question remains: "what reionized the universe?"

After reionization, gas in galaxies powered increasing rates of star formation. The star-formation rate (SFR) of the universe peaked approximately 3.5 Gyr after the Big Bang at $z \sim 2$ (Madau et al., 2014), which is often called 'cosmic noon.' About 25%

of the stellar mass observed today was formed at the peak of cosmic star formation. Galaxies in this epoch were different from galaxies in the local universe in terms of morphology, kinematics, interstellar medium (ISM) conditions, and chemical composition. We still do not fully understand these properties; we need these details to understand how these galaxies evolved into present-day galaxies.

Following cosmic noon, the cosmic star formation rate exponentially declined. Star formation is increasingly turned off (quenched) in galaxies, and they become quiescent galaxies. Several works have proposed different processes that are responsible for galaxy quenching. These processes include major mergers, quasar-mode AGN feedback, stellar winds, disk instability, and cosmic starvation. We still do not know the relative importance of these quenching mechanisms.

1.1 First billion years

The realization that the universe is ionized originated from the observation of the first quasar at $z \sim 2$ (Schmidt, 1965). Soon after, Gunn et al. (1965) found that a quasar spectrum in the blue wing of Lyα emission lacks absorption features by neutral hydrogen (Gunn-Peterson trough), implying an optical depth to Lyα scattering of only $\tau \approx 0.5$, instead of a few millions as would be expected from neutral foreground hydrogen. Two explanations were initially offered to account for the absence of the Gunn-Peterson trough: either space is empty, or hydrogen is ionized. However, since even the most under-dense voids would be enough to be optically thick to Lyα, most of hydrogen in the IGM must be ionized.

Constraints on the timing of reionization

Reionization began as early as $z \sim 11$ and ended by $z \sim 6$. There are multiple pathways to constrain the time and duration of reionization. The constraint on the beginning of reionization comes from the Planck satellite's polarization power spectrum of the cosmic microwave background (CMB). Free electrons in the ionized universe Thompson scattered quadrupole anisotropic CMB into linearly polarized CMB. Because electrons that were in causal contact with each other scattered the same anisotropic CMB sky, the observed CMB polarization is expected to fluctuate at the scale of the horizon length at reionization. This led to the "polarization bump" on ~ 10 degree scales in the polarization power spectrum. Combining this bump characteristic with the overall suppression in temperature power spectra, the latest results from Planck determine the reionization optical depth that corresponds to the redshift of instantaneous reionization at $z = 8.8^{+1.7}_{-1.4}$ (Planck Collaboration, Ade,

et al., 2016).

The constraint on the end of the reionization comes from the discovery of the Gunn-Peterson trough in a quasar spectrum at $z = 6.28$ (R. H. Becker et al., 2001). This is because even a tiny fraction of neutral hydrogen of $x_{\rm HI} \sim 10^{-4}$ can make a complete Gunn-Peterson trough in a quasar spectrum. The discovery and the presence of transmission in the Lyα forest in quasars at $z < 6$ therefore suggest that the universe must have completed its reionization by $z \sim 6$ (Fan et al., 2006).

Surveys of high-redshift Lyα-emitting galaxies (LAEs) and near-zone quasar absorption spectroscopy provide a detailed history of reionization. When the universe was partially ionized, the neutral gas in the IGM attenuates Lyα photons emitted from these high-redshift galaxies. Therefore, if we know their intrinsic Lyα flux, we can infer the neutral fraction of the IGM as a function of redshift from the observed Lyα flux. Alternatively, deep observations of the damping wings around the Lyα emission line in quasar spectra can provide additional constraints on the neutral fraction in the proximity of the quasar (Miralda-Escudé, 1998). With related assumptions on quasar and galaxy modeling, these two independent methods yield consistent constraints on the IGM neutral fraction, indicating that reionization was rather rapid. The average IGM neutral hydrogen fraction is $> 76\%$ at $z \sim 8$ (Mason, Fontana, et al., 2019), $\sim 50\%$ at $z \sim 7$ (Greig et al., 2017; F. B. Davies et al., 2018; Mason, Treu, et al., 2018) and quickly drops to $< 10\%$ at $z \sim 6$ (McGreer et al., 2015).

Sources of the ionizing photons

Soon after the realization that the universe is ionized, several works proposed sources of the ionizing photons. By the 1990s, observations of the Lyα forest constrained the IGM temperature to be $\sim 10^4$ K; it ruled out the original idea that IGM is collisionally ionized which required a temperature of $> 2\times10^5$ K (Weymann, 1967; Weinberg et al., 1997). Quasars were also soon ruled out as a significant source of the ionizing photons at $z \lesssim 4$ (e.g., Donahue et al., 1987; Miralda-Escude et al., 1990). Measurements of background X-ray photons emitted by both luminous and miniquasars at $z = 6 - 7$ limit their contribution to the required ionizing photons to $< 7\%$ (Dijkstra, Haiman, et al., 2004; Salvaterra et al., 2005). At $z \sim 8$, the current best estimate for the ionizing background provided by quasars is also at least an order of magnitude below the required ionizing background (Bouwens, Illingworth, Oesch, et al., 2015). All these lines of evidence indicate that star-forming galaxies

are the most important source of the ionizing photons (Kuhlen et al., 2012; Robertson et al., 2015).

Pinning down the contribution of ionizing photons from galaxies

Although star-forming galaxies are now accepted as primary agents that reionize the universe, calculating their actual contribution of ionizing photons remains an unfinished task. Directly measuring ionizing UV radiation from star-forming galaxies is impossible at high redshift due to IGM absorption. The ionizing UV radiation from star-forming galaxies is therefore calculated through integrating three quantities: the rest-frame UV luminosity density ($\rho_{\rm UV}$); an efficiency factor in converting the UV luminosity to Lyman-continuum photons ($\xi_{\rm ion}$); and the escape fraction ($f_{\rm esc}$). Assumptions and extrapolations on all of these parameters are still required for galaxies to provide enough ionizing photons for reionization.

UV luminosity density ($\rho_{\rm UV}$) is the luminosity per unit volume of (non-ionizing) ultraviolet photons at a rest wavelength of around 1500 Å. It is calculated by integrating galaxy UV luminosity over the UV luminosity functions of star-forming galaxies. Recent wide and deep surveys (such as Giavalisco et al., 2004; Grogin et al., 2011; Schenker, Robertson, et al., 2013; Bouwens, Illingworth, Oesch, et al., 2015) all indicate that the UV luminosity functions have steep faint-end slopes at $z \gtrsim 6$. This means that most ionizing photons come from faint galaxies. Currently, the best constraints on $\rho_{\rm UV}$ come from the Hubble Frontier Fields, whose limiting magnitude is $M_{\rm UV} \sim -14$. The UV luminosity density as a function of redshift shows a constant decline (lower $\rho_{\rm UV}$ at higher redshift) up to $z \sim 10$. The measured UV luminosity functions yield a Thomson scattering optical depth that is consistent with the *Planck* value only if they are extrapolated to fainter magnitudes $M > -14$, and the $f_{\rm esc}$ is assumed to be $> 14\%$ with $\log \xi_{\rm ion} \sim 25.3$ erg^{-1} Hz (Ishigaki et al., 2018).

The conversion factor ($\xi_{\rm ion}$) converts $\rho_{\rm UV}$ to ionizing photon density. It has also been measured extensively over the past few years, either via the UV continuum slope parameterization in the spectral energy distribution of the galaxies or via the modeling of Hα flux (e.g. Castellano et al., 2012; Dunlop et al., 2013; Robertson et al., 2015; Bouwens, Smit, et al., 2016). So far, observations and galaxy spectrum models generally agree on $\xi_{\rm ion}$: $\log \xi_{\rm ion} \gtrsim 25.3$ erg^{-1} Hz for galaxies at $z \lesssim 5$ (Bouwens, Smit, et al., 2016) with no dependence on UV luminosites for galaxies faintward of -19 mag (Bouwens, Illingworth, Oesch, et al., 2015; Lam et al.,

2019). The estimates of $\xi_{\rm ion}$ beyond $z \sim 5$ are rare. Based on a single galaxy at $z \sim 7$, there is a possibility that $\xi_{\rm ion}$ is increasing with redshift (e.g. Stark, Walth, et al., 2015), which will significantly help reduce the $f_{\rm esc}$ needed for star-forming galaxies to produce enough ionizing photons. Additional measurements of $\xi_{\rm ion}$ at high redshift are needed to clarify whether this hard ionizing spectrum is common in the reionization era.

Need for better constraints on the Lyman-Continuum escape fraction

Among the three parameters used to calculate the contribution of ionizing photons from star-forming galaxies, the Lyman-Continuum escape fraction ($f_{\rm esc}$) is least constrained by current observations. Generally, more than 10% of ionizing photons from hot stars must escape their host galaxies in order to reionize the universe by $z = 6$. For example, Atek et al. (2015) estimated the escape fraction needed to maintain the ionization of the IGM by $z \sim 7$ to be $f_{\rm esc}= 10 - 15\%$ assuming a standard value for $\xi_{\rm ion}$.

At $z \lesssim 3$, IGM absorption is not significant. It is still possible to directly observe Lyman continuum emitted by star-forming galaxies. However, direct measurements of $f_{\rm esc}$ in $z \lesssim 3$ galaxies consistently find that $f_{\rm esc}$ is $< 10\%$, and mostly on the order of a few percent (e.g. Boutsia et al., 2011; Siana et al., 2015; Izotov, Orlitová, et al., 2016). There is no clear correlation between $f_{\rm esc}$ and galaxy luminosity (Grazian et al., 2017). If $f_{\rm esc}$ of $z > 6$ galaxies is similar to that in lower-redshift galaxies, then star-forming galaxies could not have produced enough ionizing photons needed to reionize the universe.

There is clearly a need to constrain $f_{\rm esc}$ in galaxies at higher redshift. Since it is not possible to observe the Lyman continuum directly, one way around this is to use the possible relationship between the [O III]/[O II] (O32) ratio and $f_{\rm esc}$ found at lower redshift to estimate the $f_{\rm esc}$ at higher redshift (e.g. Izotov, Schaerer, et al., 2018; Faisst, 2016). However, recent studies increasingly find that the correlation between the O32 ratio and $f_{\rm esc}$ is questionable (e.g. Naidu et al., 2018; Bassett et al., 2019). Alternatively, Zackrisson, Inoue, et al. (2013) and Zackrisson, Binggeli, et al. (2017) proposed that H I clouds absorb and reprocess Lyman continuum, resulting in different changes in the observed UV spectra and Hβ emission lines. Therefore, a diagram between Hβ and UV spectral slope can be used to estimate $f_{\rm esc}$ in $z > 6$ galaxies. However, this method requires observations of Hβ emission which will only be available with the upcoming James Webb Space Telescope (JWST).

The first part of this thesis measures $f_{\rm esc}$ by using rest-frame UV absorption line optical depths to probe neutral hydrogen covering fractions (Heckman, Borthakur, et al., 2011; T. A. Jones, R. S. Ellis, et al., 2013). This work is published in Leethochawalit, T. A. Jones, R. S. Ellis, Stark, and Zitrin (2016), in which we use gravitational lensing to estimate $f_{\rm esc}$ in $z \sim 4-5$ galaxies. These are among the highest redshift estimates of $f_{\rm esc}$ by far.

1.2 Cosmic Noon

In the framework of standard ΛCDM cosmology, the large-scale structure of both dark and baryonic matter evolves under the influence of gravity. The initial fluctuations in the mass density of the early universe collapsed into mini-halos that later grew into present-day galaxies via accretion and mergers. Deep and wide-area surveys in the past decade have shown that galaxy assembly activity peaks at redshift $z \sim 3-1$ ($2-6$ Gyr after the Big Bang), as evidenced by peaks in SFR density (Madau et al., 2014) and quasar number density (Richards et al., 2006). One key aspect is that massive galaxies tend to form and assemble most of their mass earlier than lower mass galaxies (often called "downsizing"). The number density of massive galaxies ($M > M^*$, where M^* is the characteristic mass in the Schechter function) increases rapidly after $z = 4$ but stays almost constant after $z = 2$. In contrast, the number density of low-mass galaxies ($M < M^*$) increases most rapidly at $z = 1-2$ (e.g. Muzzin, Marchesini, et al., 2013).

With the above big picture, current studies on this redshift range focus on how these galaxies assemble and evolve. These include observations of galactic-scale outflows (e.g. Erb, Quider, et al., 2012; Sugahara et al., 2017; R. L. Davies et al., 2019), dynamical properties (e.g. Förster Schreiber, Genzel, Bouché, et al., 2009; Förster Schreiber, Renzini, et al., 2018; Gillman et al., 2019), gas mass fractions (e.g. Schinnerer et al., 2016; Tacconi et al., 2018), and stellar and nebular conditions (e.g. Steidel, Strom, et al., 2016; Maiolino and Mannucci, 2019). All these detailed observations are needed to test theories and simulations of galaxy formation and evolution, particularly on the processes of gas accretion and feedback from stars and black holes (Naab et al., 2017). The latter is the central theme for Chapters 3–5 of this thesis.

Specifically, in Chapter 3, we look at the effect of aggregated galactic feedback on the distribution of gas-phase metals and gas kinematics directly. We observe star-forming galaxies at 'cosmic noon,' the period with highest galaxy assembly

activity.

H II region as probes of galaxy evolution

Strong emission lines from HII regions are the tool most frequently used to probe the properties of star-forming galaxies at high redshift. Spectra of star-forming galaxies are essentially integrated spectra of stellar populations that are reprocessed by the gas and dust in the surrounding ISM. However, detecting the continuum in spectra of galaxies at higher redshift is usually difficult to do in large numbers. With current telescopes and instruments, detecting strong nebular emission lines that originated from hydrogen, oxygen, nitrogen, and sulfur in galaxies' H II regions is much more feasible. Combinations of these emission lines can reveal the properties of ionized gas (such as electron density, electron temperature, and gas chemical composition) and also the properties of the stars that ionize these H II regions.

Metallicity gradients and galactic feedback

The distribution of metal within a disk of a galaxy can be a powerful indicator of feedback strength. Disks of spiral galaxies generally follow a (double) exponential radial surface brightness profile (e.g. de Vaucouleurs, 1958; Pohlen et al., 2006). This led to the suggestion that most galaxies grow from inside out (Muñoz-Mateos et al., 2007; E. J. Nelson et al., 2012). In a spherical collapse scenario, virialized gas in the inner shells has lower initial angular momentum and therefore a shorter cooling time than that of the outer shells. As a result, gas in the inner part of a galaxy collapses and forms stars first (Roškar et al., 2008).

Based on the inside-out growth scenario, a steep metallicity gradient is expected if there is weak or no galactic feedback. All metals (except some primordial lithium) are produced by stars, either via stellar nucleosynthesis or supernovae (SN) nucleosynthesis. These metals are released back to the ISM by stellar winds and SN explosions. Thus, one would expect metals in disk galaxies to also build up from the inside out, unless there is some form of feedback and turbulence that redistribute the metals through the disks. Ma (2018) offered the following simple explanation. In a closed box model, gas-phase metallicity is a function of yield y (metals produced per unit of stellar mass formed) and gas fraction f_g:

$$Z_g = -y \ln f_g = -y \ln(1 - f_*) \tag{1.1}$$

where f_* is stellar fraction. In a gas-rich disk galaxy, this can be approximated as $Z_g \propto -\ln(1 - \Sigma_*/\Sigma) \sim \Sigma_*/\Sigma \sim \dot{\Sigma}_* t/\Sigma$. Since disks follow a power-law surface

density profile $\Sigma_g \propto R^{-\beta}$ and the Kennicutt-Schmidt law for star-formation rate $\dot{\Sigma}_* \propto \Sigma_g^{1.4} \propto R^{-1.4\beta}$, a negative metallicity gradient $Z_g \propto R^{-0.4\beta}$ is therefore expected.

Cosmological hydrodynamical simulations show that realistic feedback tends to produce relatively flat metallicity gradients in galaxy disks. Earlier works by Pilkington et al. (2012) and Gibson et al. (2013) show that when feedback is weak (e.g. only feedback from SN is included), the simulations always produce strong negative metallicity gradients (~ -0.3 dex/kpc) at $z \sim 2$. In contrast, when the simulations prescribe stronger SN feedback and include early feedback from massive stars, the metallcity gradients are flat (> -0.1 dex/kpc) at all redshifts. Based on a suite of high-resolution cosmological zoomed-in Feedback In Realistic Environments (FIRE) simulations which include multiphase feedback, Ma, Hopkins, Feldmann, et al. (2017) show that metallicity gradients in $z \sim 2$ galaxies are diverse, but mostly flat.

Gas kinematics and galactic feedback

Feedback influences not only metallicity gradients but also kinematic properties of the gas. Ma, Hopkins, Feldmann, et al. (2017) found that steep metallicity gradients (up to -0.2 dex per kpc) only appear in the simulated galaxies with well-ordered rotation ($V_{\rm rot}/\sigma \geq 1$), whereas kinematically perturbed galaxies ($V_{\rm rot}/\sigma \leq 1$) always have flat gradients. Strong feedback can induce outflows that destroy the gas rotation and redistribute metals on galactic scales. Indeed, Hung et al. (2019) found a correlation between velocity dispersion and star-formation rate in the FIRE-simulated star-forming galaxies across redshifts. Despite its large scatter, the correlation suggests that stellar-driven feedback can drive the typical amount of turbulence found in disks.

Need for adaptive optic-assisted observations of gravitationally-lensed systems

As mentioned above, observing metallicity gradients and gas kinematic properties can provide insight into the strength of the galaxies' feedback. However, this task requires spatially-resolved spectroscopic data. The angular diameter distance at $z \sim 2$ corresponds to a scale of 8.5 kpc/″, while the half-light radius of a typical L^* galaxy at $z \sim 2$ is ~ 2 kpc (Bouwens, Illingworth, Blakeslee, et al., 2004). Although an observation with adaptive optics (AO) on 8–10 meter telescopes can further refine the resolution, it can only marginally resolve $z \sim 2$ galaxies into a few resolution elements. Furthermore, Yuan, Kewley, and Rich (2013) found that coarse angular resolution can significantly bias the metallicity gradient measurements. To avoid

this bias, a spatial resolution that is finer than what an AO-assisted observation can provide is needed.

In Chapter 3, we take advantage of gravitational lensing in combination with AO observations to measure gas-phase metallicity gradients and kinematic properties of 19 $z \sim 2$ galaxies. The intrinsic physical resolutions in our sample are less than 1 kpc. The work is published in Leethochawalit, T. A. Jones, R. S. Ellis, Stark, Richard, et al. (2016). At the time, it provided the largest sample size of $z \sim 2$ galaxies at this resolution (for later results see Wang et al., 2017, in which high spatial resolution is reached via the combination of gravitational lensing and space-based observations).

Notes on gas-phase metallicity measurements

Gas-phase metallicity of the H II region is mostly measured in terms of oxygen abundance [O/H], which is the most abundant metal element in the universe. It can be measured using two methods: direct and strong-line methods. For the direct method, an electron temperature of the gas must first be calculated from the line ratios of ions with six electrons. For such ions, the radiative decays both from second to first excited state and from first excited to ground state are observable (e.g. [O III]λ4364 and [O III]λ5008). The ratio of these two lines in a low density environment, which is true for typical H II regions, depends only on the temperature. Once the electron temperature is known, the abundance of O III and O II can be determined from the ratios of their respective emission lines to a Balmer recombination line (usually Hβ). Finally, the total oxygen abundance in H II is the sum of the abundances of the two ionization states. However, this method is usually difficult in practice for high redshift galaxies because detecting the auroral lines requires long integration time. For example, the flux of [O III]λ4364 is typically less than a few percent of the [O III]λ5008 flux.

Alternatively, [O/H] can be estimated from the ratios of the strong emission lines using calibrations from either photoionization models or empirical relations. These strong-line calibrations include R23 ($\log[([\mathrm{OIII}]\lambda\lambda 4949, 5008 + [\mathrm{OII}]\lambda\lambda 3727, 3729)/\mathrm{H}\beta]$), N2 ($\log[[\mathrm{NII}]\lambda 6586/\mathrm{H}\alpha]$) and O3N2 ($\log[[\mathrm{OIII}]\lambda 5008/\mathrm{H}\beta] - \log[[\mathrm{NII}]\lambda 6586/\mathrm{H}\alpha]$) (e.g. Pilyugin, 2003; Pettini et al., 2004; Kewley and Ellison, 2008). This thesis measures oxygen abundances via the N2 and O3N2 calibrations.

One possible problem with using these calibrations is that they assume certain conditions of the H II region (such as the N/O ratio and the nature of ionizing

stars that results in a specific range of ionization parameters and O III/O II ratios). Recent studies indicate that these assumptions do not apply to all H II regions, especially those at high redshift. Although work on improving the precision of these calibrations is still ongoing (e.g. Steidel, Strom, et al., 2016; Strom et al., 2017), we expect that this does not significantly impact on our qualitative conclusion.

1.3 Epoch of galaxy quenching

Studies of the redshift evolution of galaxy luminosity functions find that the number density of L^* red galaxies has increased by a factor of a few since $z = 1$. The results indicate that at least half of the present-day L^* quiescent galaxies were star-forming at $z \sim 1$ (e.g. Faber, Willmer, et al., 2007; Cool et al., 2008). Some mechanisms must have quenched those blue galaxies and moved them onto the red sequence.

Quenching mechanisms

Many processes can be responsible for shutting down star formation in galaxies. They are either internal or external. The internal processes include the starvation or suppression of cold gas streams (e.g. Y. Peng et al., 2015; Birnboim et al., 2003), stellar feedback (e.g. Geach et al., 2014) and active galactic nuclei (AGN) feedback (e.g. Croton et al., 2006; Hopkins, Somerville, et al., 2006). External processes include wet mergers, galaxy harassment, strangulation, and ram-pressure stripping (e.g. van den Bosch et al., 2008).

It is not yet clear which quenching process is dominant but it likely depends on galaxy mass, redshift, and environment (e.g. Kawinwanichakij et al., 2017; Liu et al., 2018; Trussler et al., 2018). For stellar feedback, it is now recognized that feedback from massive stars are required in cosmological simulations in order to solve the over-cooling problem in galaxy formation. Without such feedback, gas would settle into disks too rapidly, resulting in disks that are more concentrated than observed spiral galaxies (e.g. Katz et al., 1991; Navarro et al., 1994). Even when feedback is included but is too weak (e.g., only including energy injection from supernovae), the simulations still formed too many stars (e.g. Abadi et al., 2003; Kereš et al., 2005).

Current cosmological simulations often include multiple forms of feedback that are strong enough to induce outflows. They are able to reproduce observed galaxy properties such as mass functions, galaxy mass ratios, and galaxy sizes. The problem is that they all use different feedback models (see Naab et al., 2017, for a review). Examples of these models include energy- and momentum-driven decoupled winds

with no AGN feedback (Davé, Katz, et al., 2013); momentum-driven winds with thermal/bubble AGN heating and radiation (Vogelsberger et al., 2014, the Illustris simulations); and stochastic thermal heating from stars and AGNs (Schaye et al., 2015, the EAGLE simulations).

Stellar Feedback-driven outflows

The list of the feedback models above suggests that there are mainly two types of stellar feedback wind models: energy- and momentum-driven winds. They are heuristic approaches indicating that winds are power-law functions of galaxy mass. They can be derived as following. Let v_w be wind velocity. If these winds are to become outflows, they must be larger than escape velocity, which is proportional to virial velocity: $v_w \propto v_{\rm vir}$. For energy-driven winds, the SN energy $\dot{E}_{\rm SN}$ is transferred to the wind energy $\dot{E_w} = \frac{1}{2}\dot{M}_w v_w^2$ in a conserved manner. The SN energy is also proportional to the star formation rate: $\dot{E}_{\rm SN} \propto \dot{M}_*$. Equating $\dot{E}_{\rm SN} = \dot{E_w}$ yields

$$\frac{\dot{M}_w}{\dot{M}_*} \propto v_{\rm vir}^{-2}. \tag{1.2}$$

The term $\frac{\dot{M}_w}{\dot{M}_*}$ is the mass outflow per star formation rate, usually defined as the mass-loading factor (η). For momentum-driven winds, the wind kinematics are governed by momentum conservation: $\dot{M}_w v_w \propto \dot{M}_*$, which simply means that

$$\eta \propto v_{\rm vir}^{-1}. \tag{1.3}$$

This suggests that the mass-loading factor is likely a function of galaxy mass. For example, we can use an empirical relation between stellar mass and central stellar velocity dispersion, e.g., $\sigma \propto M_*^{0.3}$ (Zahid, Geller, Fabricant, et al., 2016). This predicts the relations $\eta \propto M_*^{-0.3}$ and $\eta \propto M_*^{-0.6}$ for momentum- and energy-driven wind models, respectively.

Need for measurements of galactic outflows

One way to differentiate between these feedback models is to obtain better observational constraints on galactic outflows, such as velocities, mass outflow rates, their spatial extent, and the scaling of these quantities with galaxy mass. Holistic observations of galactic outflows is challenging because outflows are multi-phase, ranging from cold molecular gas to hot plasma. If the galaxies are near enough, emission lines from diffuse outflowing gases such as Hα and [N II] can be directly observed (e.g. Filippenko et al., 1992). For further galaxies, the most commonly-used method to trace outflows is the “down-the-barrel" technique, the spectroscopy

of blue-shifted interstellar absorption lines from cool, outflowing gas, seen against the galaxy's stellar continuum (e.g. Martin, 2005; Erb, Quider, et al., 2012; Rubin et al., 2014). This method can directly measure outflow velocities. However, calculating mass outflow rates is more difficult as it involves assumptions on outflow spatial profile, metallicity, and ionization fractions. Nonetheless, this method has so far been providing most of the numbers for the mass-loading factors in star-forming galaxies (e.g., S. F. Newman, Genzel, Förster-Schreiber, et al., 2012; Chisholm et al., 2017; T. Jones et al., 2018).

Alternatively, outflows can be measured from their imprint on galaxy metal abundances. It is crucial to explain the mass-metallicity relation (MZR) of the galaxies, in which both gas and stellar metallicities correlate with galaxy mass. Galaxies with higher mass have higher metallicities than galaxies with lower mass. Conceptually, this is because massive galaxies have deeper gravitational potentials and are expected to retain more of their SN ejecta than dwarf galaxies (e.g. Larson, 1974; Dekel et al., 1986).

In the last two chapters of this thesis, we spectroscopically study quiescent galaxies at $z < 1$, the period during which star formation rate density of the universe declines. The goal is to measure the galaxies' past outflow history based on their surviving chemical composition. The results provide alternative constraints on galactic outflows in terms of mass-loading factors. The works are published in Leethochawalit, Kirby, et al. (2018) and Leethochawalit et al. (submitted to ApJ).

Notes on stellar metallicity measurements

Stellar metallicities are measured from stellar light. If the galaxy is close enough, i.e., in the Local Group, spectra of individual stars can be obtained. Its metalicity is therefore calculated from metallicities of individual stars, either spectroscopically or photometrically. However, most galaxies are too far to be resolved into individual stars. As a result, their metallicities have to be measured from integrated spectra of their stellar population.

Historically, integrated-light stellar metallicities have been measured through the analysis of spectral indices, i.e., equivalent widths (EWs) of absorption lines or molecular bands (e.g., Worthey et al., 1994, the Lick indices). A set of measured EWs is converted to metallicities based on calibrations from stellar population models. This method is perceived to be convenient because it dissects galaxy spectra into features of individual elements, is insensitive to dust attenuation, and

requires less computing capabilities. However, in practice, these absorption lines are not truly distinct because their EWs can only be measured based on the observed psuedo-continuum levels, which depend on abundances of multiple elements (see Conroy, Graves, et al., 2014, for a review).

A more recent method to derive detailed abundances of a galaxy is via full-spectrum fitting. Its main advantage is that it uses all information from the observed spectra, rather than regions of the selected absorption lines. Therefore, with the same S/N, the full-spectrum fitting technique results in smaller uncertainties in the measured parameters (Conroy, Villaume, et al., 2018). In the past decade, efforts have been put into improving stellar population models. Currently, there are a number of spectral fitting algorithms available, such as pPXF (Cappellari, 2017), STARLIGHT (Cid Fernandes et al., 2005), STECKMAP (Ocvirk et al., 2006), FIREFLY (Wilkinson et al., 2017) and `alf` (Conroy, Villaume, et al., 2018).

Regardless, the two techniques used to derive stellar metallicities both rely on stellar population synthesis (SPS). The base model of every SPS is simple stellar populations (SSPs), which describe spectra of a coeval stellar population (single age) at a single metallicity. An SSP spectrum is created by integrating spectra of all stars of the same age and metallicity (i.e., stars that are on the same stellar isochrone), mass weighted by an initial mass function (Φ, IMF). The formula is

$$f_{\rm SSP}(t, Z) = \int f_*(T_{\rm eff}(M), \log g(M)|t, Z)\Phi(M)dM \tag{1.4}$$

where f_* is stellar spectra in a chosen stellar library with effective temperature $T_{\rm eff}$ and surface gravity $\log g$ determined by the isochrone of age t and metallicity Z. A more complex spectral model can be created by combining the SSP spectra according to the galaxy's star formation history (SFH). It can also be later reddened by dust and added with emission lines from gas. However, to actually constrain SFH, spectra with high S/N are generally required.

All stellar metallicities in this thesis are measured via the full-spectrum fitting technique. We fit the observed spectra with SSP spectral models from Conroy, Gunn, and M. White (2009) and Conroy, Villaume, et al. (2018).

1.4 Goals of this thesis

This thesis uses spectroscopy of galaxies to better characterize roles of galactic feedback on galaxy evolution. We observe a wide range of galaxy types over cosmic time: clumpy star-forming galaxies in the early universe, primitive star-forming

disks during their emergence, and quiescent galaxies in the second half of cosmic time. The specific goals are as follow.

1. In Chapter 2, we use Keck/DEIMOS deep optical spectroscopy to probe the rest-frame UV spectra of high-redshift star-forming galaxies. The main goal is to measure escape fractions $f_{\rm esc}$ at the epoch as close to reionization as possible. The quantification of escape fractions at this redshift range is needed to confirm whether star-forming galaxies were the primary agent in reionizing the universe. In total, we observe four $z \sim 4-5$ gravitationally-lensed galaxies and measure their $f_{\rm esc}$ via the measurements of their covering factions. As by-products, we are able to probe the strengths and roles of star formation feedback in the early universe.

2. In Chapter 3, we use Keck/OSIRIS near-infrared, spatially-resolved spectroscopy to probe the strength of stellar feedback in primitive galaxy disks at the peak of star formation. Specifically, we observe 11 gravitationally-lensed galaxies and measure their gas-phase metallicity gradients and kinematic properties. With the boost in both S/N and angular resolution from gravitational lensing, we are able to measure metallicity gradients and model rotation at the sub-kpc scales.

3. In Chapters 4 and 5, we use Keck/DEIMOS and LRIS optical spectroscopy to probe the stellar populations of quiescent galaxies at $z < 1$. We measure stellar metallicities (both [Fe/H] and [Mg/H] abundances) via spectral synthesis. The goal is to measure the mass–metallicity relations (MZRs) at three redshifts: $z \sim 0, 0.39$, and 0.55. Lastly, we apply galactic chemical evolution models to the MZRs in order to archaeologically quantify galactic outflows in terms of mass-loading factors.

Chapter 2

ABSORPTION LINE SPECTROSCOPY OF GRAVITATIONALLY-LENSED GALAXIES: FURTHER CONSTRAINTS ON THE ESCAPE FRACTION OF IONIZING PHOTONS AT HIGH REDSHIFT

Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. (2016), ApJ 831, 152, 152

Abstract

The fraction of ionizing photons escaping from high-redshift star-forming galaxies remains a key obstacle in evaluating whether galaxies were the primary agents of cosmic reionization. We previously proposed using the covering fraction of low-ionization gas, measured via deep absorption line spectroscopy, as a proxy. We now present a significant update, sampling seven gravitationally-lensed sources at $4 < z < 5$. We show that the absorbing gas in our sources is spatially inhomogeneous with a median covering fraction of 66%. Correcting for reddening according to a dust-in-cloud model, this implies an estimated absolute escape fraction of $\simeq 19 \pm 6\%$. With possible biases and uncertainties, collectively we find the average escape fraction could be reduced to no less than 11%, excluding the effect of spatial variations. For one of our lensed sources, we have sufficient signal/noise to demonstrate the presence of such spatial variations and scatter in its dependence on the Lyα equivalent width, consistent with recent simulations. If this source is typical, our lower limit to the escape fraction could be reduced by a further factor $\simeq$2. Across our sample, we find a modest anti-correlation between the inferred escape fraction and the local star formation rate, consistent with a time delay between a burst and leaking Lyman continuum photons. Our analysis demonstrates considerable variations in the escape fraction consistent with being governed by the small-scale behavior of star-forming regions, whose activities fluctuate over short timescales. This supports the suggestion that the escape fraction may increase toward the reionization era when star formation becomes more energetic and burst-like.

2.1 Introduction

Considerable observational progress is being made in charting when cosmic reionization occurred. The most recent values of $\tau = 0.058 \pm 0.012$, the optical depth of electron scattering of cosmic microwave background (CMB) photons, reported by the Planck consortium (Planck Collaboration, Ade, et al., 2016; Planck Collaboration, Adam, et al., 2016), implies the average redshift at which reionization occurred lies between $z = 7.8$ and 8.8. This represents a significant revision of the picture inferred from earlier CMB data and indicates reionization was both more rapid and occurred more recently than earlier imagined. Combining the CMB data with line of sight measures of the opacity of the intergalactic medium (IGM) obtained via spectroscopy of high redshift quasars and gamma ray bursts (Chornock et al., 2014) and the fraction of Lyman break galaxies whose Lyα emission is detectable (Schenker, R. S. Ellis, et al., 2014), indicates reionization began shortly before $z \simeq 10$ and was largely complete by $z \simeq 6$.

Late reionization is particularly important in strengthening the view that star-forming galaxies were the primary ionizing agent. The abundance and luminosity distribution of $6 < z < 10$ galaxies is increasingly well-determined from deep imaging in blank and cluster lens fields using Hubble Space Telescope's (HST) near-infrared camera WFC3/IR(Oesch et al., 2015; Finkelstein et al., 2015; McLeod et al., 2016; Atek et al., 2015). Robertson et al. (2015) present a model showing that, depending on their ionizing output, the demographics of the $z > 6$ galaxy population can match the Planck τ values by a redshift $z \simeq 10$. There are two key uncertainties in the ionizing output. The first is an efficiency factor, $\xi_{\rm ion}$, necessary to convert the (observed) UV luminosity to the production rate of Lyman-continuum photons, which depends on the nature of the stellar population and the presence or otherwise of AGN. Reasonable progress has been made in estimating typical $\xi_{\rm ion}$ values for high redshift galaxies from recombination or other emission line measures (Bouwens, Illingworth, Oesch, et al., 2015; Stark, R. S. Ellis, Charlot, et al., 2017). The second uncertainty is the fraction, $f_{\rm esc}$, of the ionizing photons that escape into the IGM due to the porosity of neutral hydrogen in a typical galaxy. For reasonable values of $\xi_{\rm ion}$, Robertson et al. (2015), Bouwens, Illingworth, Oesch, et al. (2015) and Khaire et al. (2016) require an absolute escape fraction, $f_{\rm esc,abs}$[1], of up to 20% to account for the Planck value of τ whereas Atek et al. (2015) and Mitra et al. (2015) estimated $f_{\rm esc,abs} \approx 10 - 15\%$.

[1] See Section 2.4 for the definition of $f_{\rm esc,abs}$ and $f_{\rm esc,rel}$.

The most direct way to measure the escape fraction is to detect the leaking Lyman continuum (LyC) photons through spectroscopy or appropriate rest-frame imaging. In the local universe, non-zero spectroscopic LyC measures are only available for six galaxies. Most have $f_{\rm esc,abs} < 5\%$ (Leitet, Bergvall, Piskunov, et al., 2011; Leitet, Bergvall, Hayes, et al., 2013; Borthakur et al., 2014; Leitherer et al., 2016), while one has $f_{\rm esc,abs} \sim 8\%$ (Izotov, Orlitová, et al., 2016). Several surveys have attempted direct detection at $z \sim 2-3$ with a similarly low success rate - a few detections out of hundreds examined (e.g. Vanzella, de Barros, Castellano, et al., 2015; Vanzella, de Barros, Vasei, et al., 2016; Mostardi et al., 2015). Even when significant detections are apparent, they usually reflect contaminating radiation from foreground galaxies that can easily camouflage as substructures of the clumpy high-redshift galaxies(e.g. Siana et al., 2015; Mostardi et al., 2015). So far, only three convincing detections have been reported: $f_{\rm esc,rel} \simeq 42\%$, $f_{\rm esc,rel} > 75\%$ and $f_{\rm esc,rel} > 50\%$[1] (Mostardi et al., 2015; Vanzella, de Barros, Vasei, et al., 2016; Shapley, Steidel, Strom, et al., 2016). With an assumption of some attenuation, Mostardi et al. (2015) translates the measured relative escape fraction to $f_{\rm esc,abs} = 14\% - 19\%$. Though higher than local measurements, the measured escape fractions at $z \sim 3-4$ are still relatively lower than expected considering that only the high tail of the $f_{\rm esc}$distribution can be readily detected using these techniques. The relatively rare LyC leakers at $z \lesssim 3$ tend to be compact, less massive, efficiently star-forming systems with high [O III]/[O II] emission line ratios (Leitet, Bergvall, Piskunov, et al., 2011; Alexandroff et al., 2015; Izotov, Orlitová, et al., 2016). Nakajima et al. (2014) present a model whereby the high [O III]/[O II] line ratio seen in many intense Lyα emitters can be understood if the H II regions are density-, rather than ionization-, bound. In this case, although a rare population locally, if such [O III]-intense systems are more common at high redshift(e.g. Roberts-Borsani et al., 2016; Smit et al., 2014), conceivably the mean escape fraction in the reionization era is significant (e.g. Faisst, 2016).

Numerical simulations show mixed results regarding escape fraction and its evolution with redshift. Hydrodynamic simulations yield escape fractions ranging from smaller than a few percent (e.g. Wise et al., 2014; Paardekooper et al., 2015; Ma, Hopkins, Kasen, et al., 2016) to $\sim 10\%$ (e.g. Kimm et al., 2014) in galaxies with halo masses $M \gtrsim 10^9 M_\odot$ with a strong anti-correlation between escape fraction and halo mass. Ma, Hopkins, Kasen, et al. (2016) show recently that inclusion of binary star evolution can boost the escape fraction at $z \sim 6$ to be as high as $\sim 20\%$. Most simulations find little or no evolution in the escape fraction but rather a temporary fluctuation with different delay times after bursts of star formation (e.g. Kimm et al.,

2014).

While observational progress will eventually be made in securing robust measures of the escape fractions for large samples at $z \lesssim 3$, the above discussion illustrates how evolution in the galaxy population necessitates measures of $f_{\rm esc}$ for sources in the reionization era. Unfortunately, the direct methods discussed above are not practical for galaxies with $z \gtrsim 3$ due to the increased IGM absorption. Only a modest amount of foreground neutral hydrogen can completely absorb the LyC radiation. Therefore, even with the full armory of the James Webb Space Telescope (JWST) and next generation large ground-based telescopes, indirect methods are the only route to estimating $f_{\rm esc}$ and hence understanding the role of galaxies in this final piece of cosmic history. Zackrisson, Inoue, et al. (2013) have proposed a method for estimating the LyC leakage using a combination of the UV spectral energy distribution (which constrains the stellar population) and the strength of a Balmer line such as Hβ through recombination theory. For $z > 6$ galaxies, this requires spectroscopic access to wavelengths beyond 2 μm and thus must await JWST.

A more immediate indirect method is to constrain the *c*overing fraction f_{cov} of low ionization gas from absorption line spectroscopy. As discussed in Section 2.4, this provides a valuable proxy for the escape fraction. The technique was first applied locally with promising results. Heckman, Borthakur, et al. (2011) studied 26 star-forming galaxies correlating the residual absorption signals they found from low ionization species with FUSE spectroscopy below the Lyman limit and the strength of recombination lines. In our earlier work (T. Jones et al., 2012) we first applied the absorption line method to a stacked spectrum derived from 81 $z \simeq$4 Lyman break galaxies (LBGs). In comparison with a similar stack at $z \simeq 3$ (Shapley, Steidel, Pettini, et al., 2003), the covering fraction was reduced implying an increased porosity of neutral gas and a higher $f_{\rm esc}$. A further comprehensive analysis of stacked spectra in the context of the escape fraction for 121 $z \simeq$3 galaxies at $z \simeq$3 has recently been presented by Reddy et al. (2016).

However, our experience with stacked spectra has revealed limitations in their utility for the present purpose. At the spectral resolution available in our large scale redshift survey at $z \simeq 4-5$(T. Jones et al., 2012), the weakened absorption found at $z \simeq$4-5 could equally reflect evolution in the outflow kinematics. By its very nature as a composite, the stacked spectrum must average over of a range of absorption line kinematic profiles, confusing any interpretation of the spectra in terms of opacity alone. Clearly it is preferable to study the joint kinematics and absorption line depths

in suitably bright individual galaxies. Although a challenging task, in T. A. Jones, R. S. Ellis, et al. (2013) we secured higher resolution absorption line spectra of 3 *g*ravitationally-lensed $z \simeq 4$ galaxies and, taking into account different kinematic profiles, confirmed the original redshift dependence mentioned in T. Jones et al. (2012) We also found a lower covering fraction for those galaxies with intense Lyα emission. Although one expects $f_{\rm esc} < 1 - f_{\rm cov}$ and thus the method can strictly only provide an upper limit (T. A. Jones, R. S. Ellis, et al., 2013; Vasei et al., 2016), our goal here is to enlarge the sample of lensed galaxies, both in size and redshift range, so we can strengthen evidence for an evolutionary trend as well as improve our understanding of the behavior of $f_{\rm esc}$ with other physical properties of high redshift galaxies.

In this work, we analyze high signal to noise absorption line spectra for a further 4 lensed high redshift galaxies, more than doubling the sample in T. A. Jones, R. S. Ellis, et al. (2013) and extending the redshift range to $z \simeq 5$. This allows us to better examine the evolutionary trend in $f_{\rm cov}$ and correlations with other parameters. Moreover, by studying lensed galaxies with strong gravitational magnification, we demonstrate that resolved spectroscopy allows us to examine how the covering fraction varies *a*cross a galaxy as a function of its kinematics and spatially-dependent star-formation rate.

A plan of this chapter follows. We present our new data in Section 2.2. The basic properties of our sample are then determined in Section 2.3, including the lens models essential for recovering the intrinsic star formation rates and luminosities. We introduce two methods for deriving the escape fractions in Section 2.4 and examine the trends derived for the total sample (including that in T. A. Jones, R. S. Ellis, et al. 2013) with redshift and the Lyα equivalent width. We also discuss the spatially-resolved data for the $z = 4.92$ galaxy lensed by the cluster MS1358+62. We summarize our results in Section 2.5.

2.2 Data

For the present study, we selected four gravitationally-lensed galaxies from the literature using the following criteria: (i) spectroscopic redshift $z > 3$, (ii) total apparent magnitude $I_{AB} < 23$ and (iii) the availability of a reasonable lens model to enable the magnification to be determined. In conjunction with the earlier study by T. A. Jones, R. S. Ellis, et al. (2013), this provides a sample of 7 lensed galaxies in the redshift range $4 < z < 5$.

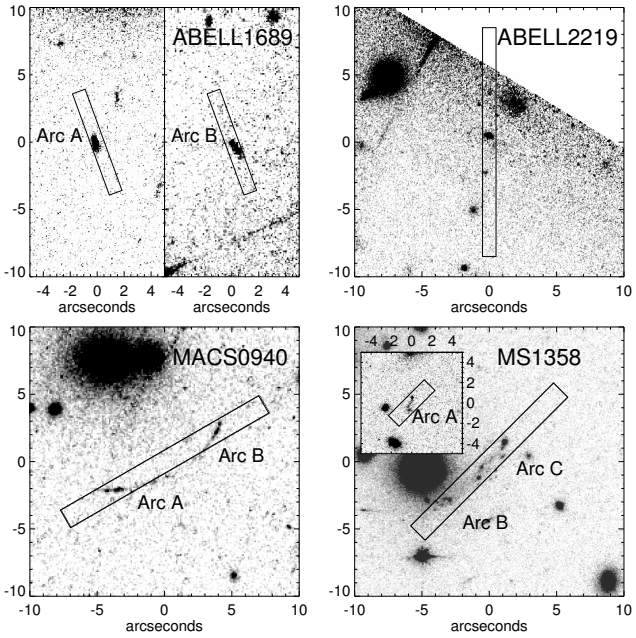

Figure 2.1: Spectroscopic configuration for the four new arcs. The DEIMOS slit positions are overlaid on HST images for (left to right) Abell1689, Abell2219, MACS0940 and MS1358. North is up and east is to the left. The instrument/filter are ACS/F850LP for Abell 1689, Abell 2219 and MS1358, and WFPC2/F606W for MACS0940.

We observed each of the four new galaxies with the DEIMOS multi-slit spectrograph mounted on the Keck II telescope. Following the approach of T. A. Jones, R. S. Ellis, et al. (2013), we used a 1200 line mm^{-1} grating yielding a spectral resolution of 1.8 Å determined from atmospheric emission lines which corresponds to a velocity resolution of 70 km sec^{-1} (see details of the observations record in Table 2.1). This grating choice allow us to sample rest-frame wavelengths from 1200 to 1650 Å ensuring good coverage of Lyα, Si IIλ1260, O Iλ1302+Si IIλ1304, C IIλ1334, and Si IIλ1527. For multiply-imaged galaxies, we arranged the multi-slits to cover as many images as practical. The final slit placements are overlaid on HST images of each target in Figure 2.1.

We reduced the DEIMOS spectra using code associated with the SPEC2D pipeline. Because our galaxies are lensed with extended non-Gaussian profiles, we used a boxcar extraction with specified full width half maxima (FWHM) to extract the 1D spectra from the 2D data. Each spectrum was corrected for telluric absorption line using spectra of standard stars and smoothed with the spectral resolution at each wavelength measured from atmospheric emission lines. For absorption line measurements, we spline fitted the continuum and normallized each spectrum. If

Table 2.1: Observation Data

ID	Arc	RA	DEC	$z_{\rm ISM}$	$z_{{\rm Ly}\alpha}$	Date	$t_{\rm exp}$(hr)
MACS 0940	ArcAB	09:40:53	+07:44:18	4.031 ± 0.001	4.034	03/15	5.5
Abell 2219		16:40:16	+46:43:59	4.450 ± 0.002	4.451	05/15	6.2
Abell 1689	Arc A	13:11:25	-01:20:52	4.868 ± 0.001	4.875	05/15	4.9
	Arc B	13:11:31	-01:20:14				
MS 1358	ArcA	13:59:55	+62:31:04	4.928 ± 0.002	4.927	03/15&02/16	7.5
	ArcBC	13:59:49	+62:30:48				
Objects published in T. A. Jones, R. S. Ellis, et al. (2013)							
Abell2390_H3a		21:53:34	+17:42:03	4.043 ± 0.002		10/11,06/12	10.8
Abell 2390_H5b		21:53:35	+17:41:33	4.0448 ± 0.0003		10/11,06/12	10.8
1621+0607		16:21:33	+06:07:05	4.1278 ± 0.0004		10/11,06/12	5.7

the galaxy was observed with multiple masks (in the case of MS1358) or had multiple images (MACS0940), we combined the continuum-normallized 1D-spectra using a weighted arithmetic mean. For MS1358, we excluded Arc B in the final combination both to prevent contamination from a foreground lensing galaxy and because some sub-components in the lensed image lay outside the slit. For star formation rate measurements, we calibrated each telluric-corrected spectrum using standard stars observed at the end of each night for an overall throughput correction and flux calibrated using HST photometry in that imaging band which overlaps our DEIMOS wavelength coverage. The spectra of each galaxy were finally combined using weighted arithmetic mean. The final spectra are shown in Figure 2.2; the typical signal-to-noise is ~ 10 per spectral resolution element.

2.3 Derived Properties

We first derive the basic physical properties of our sample including the magnification for each lensed galaxy determined from a cluster mass model. These are summarized in Table 2.

Systemic Redshift

The systemic redshift is an important measure since it enables us to examine the spectra for the presence of outflowing gas. Accurate systemic redshifts are usually obtained from stellar absorption lines or the nebular emission lines arising from H II regions. For our sample, a redshift from nebular [O II] 3727 Å emission is only available for MS1358 (Swinbank, Webb, et al., 2009). However, since our spectra primarily target low-ionization absorption lines and Lyα emission, it is more practical to use these to estimate the systemic redshift (which is in reasonable

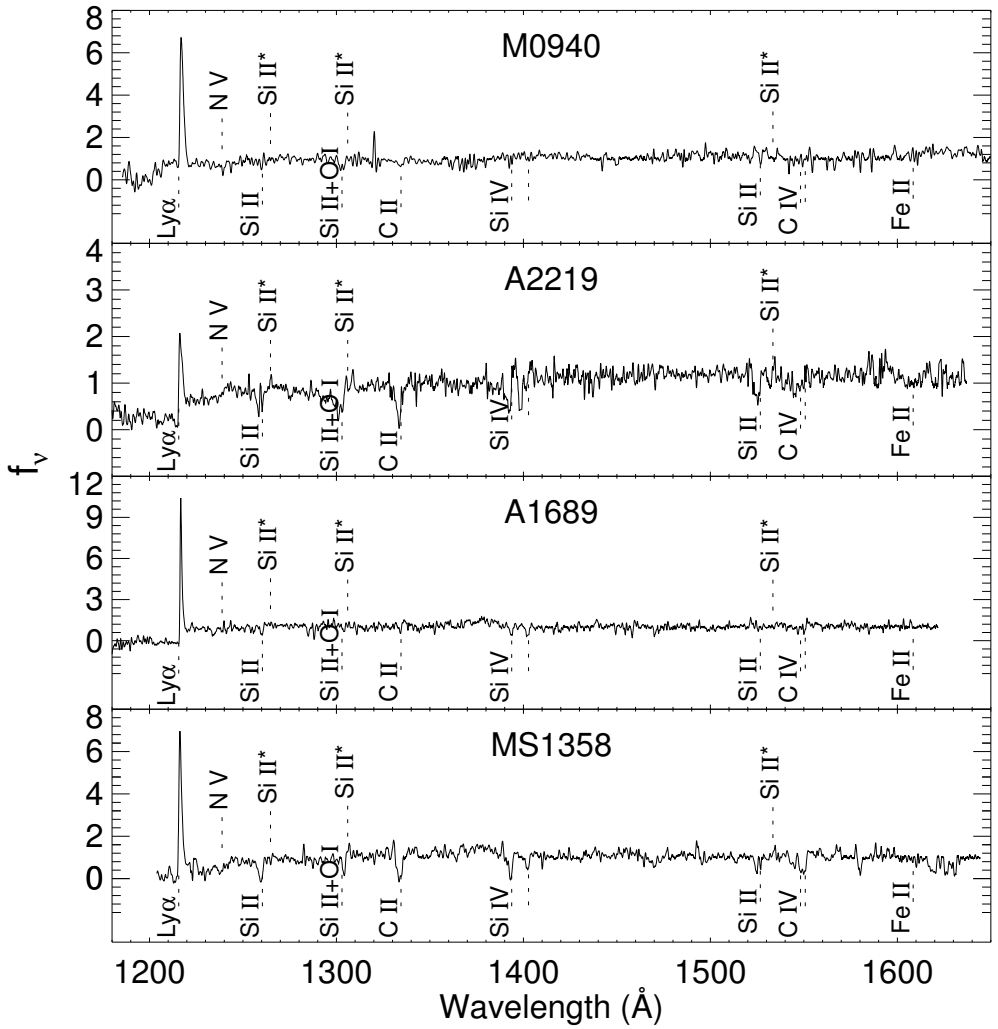

Figure 2.2: Combined absorption line spectra for the four arcs with the low ionization lines marked.

agreement with that derived from nebular emission for MS1358). T. Jones et al. (2012) determined an offset of Δv_{IS}=190 km s^{-1} between the systemic redshift (z_{sys}) determined from the C IIIλ1176 stellar photospheric line to that measuring using the low-ionization absorption lines (z_{IS}) using a composite spectrum of 81 LBGs at $z \sim 4$. Following T. A. Jones, R. S. Ellis, et al. (2013), we apply these velocity offsets to the average weighted centroid of low-ionization absorption lines to determine the systemic redshift of each galaxy. We used all prominent low-ionization absorption lines available namely Si IIλ1260, O Iλ1302, Si IIλ1304, C IIλ1334, and Si IIλ1527. The uncertainty is a square sum of the standard deviation of redshifts measured from different absorption lines and the rms scatter $\sim$ 125 km s^{-1} between redshifts derived from absorption lines and those from nebular emission (Steidel, Erb, et al.,

2010). Together with the redshift of Lyα, the results are listed in Table 2.1.

Star Formation Rates

We measure the star formation rates (SFR) of our lensed targets using the extinction-corrected 1600 Å flux from the observed DEIMOS spectra. To account for dust extinction, we measure the UV continuum slope(β)[2] of the spectra via a linear fit to magnitude as a function of logarithmic wavelength over 1250-1650 Å in rest frame. It is not possible to estimate the UV slope using photometric data for 3 of our recently-observed sample due to the absence of one or more of the key photometric bands. However, for MS1358 we can compare our spectroscopically-measured slope with that derived from available HST/ACS F850LP and HST/WFC3 F110W data. We find reasonable agreement given the uncertainties arising from the faint and extended nature of this object. Our UV slopes range from $\beta \sim -1.4$ to -2.8 in good agreement with those determined for large samples of LBGs at similar redshifts (e.g. Bouwens, Illingworth, Franx, et al., 2009; Castellano et al., 2012). The Calzetti et al. (2000) extinction law is applied to translate the UV slope to the extinction at 1600 Å. The SFR is then calculated from the extinction-corrected 1600 Å flux according to the relation given by Kennicutt (1998), in which a Salpeter (1955) initial mass function was assumed. Finally, we used the magnification from the lens models described in Section 2.3 to derive the intrinsic SFR. For the three lensed galaxies from T. A. Jones, R. S. Ellis, et al. (2013), we adopt a similar procedure except the UV slopes were determined from the appropriate rest-frame UV HST photometry. For Abell 2390_H3a, the UV slope is calculated from HST/WFPC2 F814W and Keck/NIRC H band photometry reported in A. J. Bunker et al. (2000). For Abell 2390_H5b and J1621+0647, the UV slopes are calculated from HST/WFPC2 F814W, HST/WFC3 F125W, and F110W images.

For the arc in the M1358+62 cluster, we can compare our derived SFR to that estimated from other methods in the literature. For the north-west arc (Arc C in Figure 2.1), we measure $\mathrm{SFR_{UV}} = 594 \pm 202\ \mathrm{M_\odot yr^{-1}}$ (uncorrected for lensing). The result is consistent with that derived from an [OII] emission line flux with no reddening correction at $\mathrm{SFR_{O[II]}} = 525 \pm 55\ \mathrm{M_\odot yr^{-1}}$ (Swinbank, Webb, et al., 2009). Similarily, the unlensed SFRs for Abell 2390_H3a and Abell 2390_H5b of 95 ± 52 and $13 \pm 5 M_\odot\ \mathrm{yr^{-1}}$ are consistent with estimates reported in Pelló et al. (1999) using SED-model fittings of ~ 60 and $\sim 10 M_\odot\ \mathrm{yr^{-1}}$ respectively.

[2]Ultraviolet continua are usually parametrized via a power-low index, β, where $f_\lambda \propto \lambda^\beta$ (Meurer et al., 1999)

Table 2.2: Extinction and Star Formation Rates

Cluster	Arc	Magnification (μ)	UV slope (β)	SFR ($M_\odot$ yr^{-1})	$\overline{\text{SFR}}$ ($M_\odot$ yr^{-1})
MACS 0940	arcA	182 ± 411	-1.49 ± 0.11	1 ± 3	2 ± 3
	arcB	38 ± 19	-1.35 ± 0.26	9 ± 10	
Abell 2219		1.6 ± 0.3	-1.76 ± 0.09	89 ± 17	
Abell 1689	arcA	5.3 ± 1.1	-1.76 ± 0.07	46 ± 9	47 ± 9
	arcB	2.2 ± 0.4	-1.92 ± 0.21	44 ± 45	
MS 1358	arcA	2.9 ± 0.6	-1.60 ± 0.11	74 ± 14	73 ± 14
	arcC	12.1 ± 6.1	-1.67 ± 0.04	66 ± 33	
Abell 2390_H3a		19 ± 10	-1.75 ± 0.10	95 ± 52	
Abell 2390_H5b		6.5 ± 1.3	-2.81 ± 0.15	13 ± 5	
1621+0607		38 ± 19	-1.9 ± 0.2	31 ± 20	
Clumps in MS1358					
MS1358	C0	10.0 ± 2.0	-1.74 ± 0.07	33 ± 6	
MS1358	C1,2,3	11.5 ± 5.8	-1.89 ± 0.08	15 ± 7	
MS1358	C4	13.1 ± 6.6	-1.99 ± 0.09	6 ± 3	

Lens Models and Magnifications

Gravitational lensing enables us to directly measure covering fractions for galaxies at high redshift without the need for lens models since the magnification involved has no wavelength dependence. However, to derive the intrinsic properties such as the star formation rate, an accurate mass model for the lensing cluster is required. For each galaxy cluster we constructed a mass model using the Light-Traces-Mass (LTM) method (full details in Zitrin et al. 2015a and Zitrin, Broadhurst, Umetsu, et al. 2009). The model consists of a galaxy component and a smooth dark matter (DM) component. The galaxy component is a superposition of all cluster member galaxies, where the model assumes the mass of each galaxy scales with its luminosity and can be described via an elliptical shaped power law profile. A core is often also introduced, especially for the brightest central galaxies (BCGs). The mass-to-light ratio is fixed for all cluster members except the BCG. The DM component is a smoothed version of the galaxy component, and the two components are added with a relative weight optimized in the minimization procedure. In addition, a two-component external shear is added to allow for further flexibility. We use a Markov Chain Monte Carlo (MCMC) to construct the mass model by minimizing differences between the predicted and actual position of multiple images. The multiple images and their spectroscopic redshifts were obtained from the literature i.e.,Swinbank, Webb, et al. (2009), Zitrin, Broadhurst, Coe, et al. (2011), Smith et al. (2005),

Gladders et al. (2002), and Rzepecki et al. (2007).

Using the mass model, we transfer the rest-frame UV image into the source plane and determine the magnification μ from the associated flux ratio. Uncertainties arise both from the precision of the model, which can be estimated from the predicted versus observed position of multiple images, and systematic effects associated with the LTM model compared, for example with a model based on analytic profiles (see Zitrin et al. (2015a)). We adopt an uncertainty of 20% when the magnification $\mu < 10$ and 50% if $\mu > 10$. The magnification $\mu = 12.1 \pm 6.1$ derived for MS1358 is consistent with the value of $\mu \simeq 12.5$ derived by Swinbank, Webb, et al. (2009) offering a valuable confirmation of our methods.

2.4 Covering and Escape Fractions

Definitions

The goal of this paper is to interpret the covering fraction of low ionization gas in terms of new constraints on the escape fraction of ionizing photons in an enlarged sample of $4 < z < 5$ galaxies. We will introduce two methods for making the connection between the covering and escape fraction. However, it is important to introduce some definitions first.

The most important measure of the escape fraction is that used in calculating whether galaxies are capable of driving cosmic reionization. This is the *a*bsolute escape fraction defined as:

$$f_{\rm esc,abs} \equiv \frac{F_{\rm LyC,out}}{F_{\rm LyC,int}}$$

that is the ratio between the Lyman continuum flux density escaping the galaxy, usually measured at 900 Å, to the internal stellar flux density at the same wavelength. Most of the escape fractions reported in numerical simulations and theoretical calculations refer to this absolute value. However, observationally it is very difficult to estimate the internal stellar flux density at 900 Å since it requires modeling the spectral energy distribution(SED) beyond the range of direct observation. Instead, a more accessible estimate is the *r*elative escape fraction defined as:

$$f_{\rm esc,rel} \equiv \frac{(f_{\rm LyC}/f_{1500})_{\rm out}}{(f_{\rm LyC}/f_{1500})_{\rm int}} = \frac{f_{\rm esc,abs}}{10^{-0.4A_{1500}}}$$

Since $f_{\mathrm{LyC,out}}$ can be additionally absorbed by neutral hydrogen in the IGM along the line of sight, we can write $f_{\mathrm{esc,rel}}$ as a function of the observed Lyman continuum flux($f_{\mathrm{LyC,obs}}$) as

$$f_{\mathrm{esc,rel}} = \frac{(f_{\mathrm{LyC}}/f_{1500})_{\mathrm{obs}}}{(f_{\mathrm{LyC}}/f_{1500})_{\mathrm{int}}} e^{\tau_{\mathrm{HI,IGM}}}$$

Notwithstanding the fact that the relative escape fraction can be obtained from good observations, modeling of the SEDs is still required to determine the intrinsic ratio of the 900 Å and 1500 Å luminosities, which results in additional uncertainties.

Method 1: Low-ionization Covering Fraction

We can estimate the covering fraction of neutral hydrogen in a galaxy by measuring the amount of non-ionizing UV radiation absorbed by metals following the approach described in Heckman, Borthakur, et al. (2011) and T. A. Jones, R. S. Ellis, et al. (2013). Realistically, we must assume that galaxies are covered partially by H I regions with different velocities relative to the systemic value. Neutral clouds of a particular velocity will have a covering fraction of $f_{\mathrm{cov}}(v)$. Photons of each low-ionization absorption line, when shifted to match the cloud velocity, will be able to escape if they travel through holes in the neutral gas or will otherwise be attenuated. The underlying principle is that low ionization gas represents a faithful tracer of neutral hydrogen, an assumption discussed in detail by T. A. Jones, R. S. Ellis, et al. (2013), Henry et al. (2015) and Vasei et al. (2016). Trainor et al. (2015) find low ionization absorption line widths correlate with the inferred escape of Lyman α photons, and Rivera-Thorsen et al. (2015) find a strong correlation with 21 cm line widths which directly trace the neutral hydrogen. However, Reddy et al. (2016) claim, from their recent $z \simeq 3$ sample, that the outflowing gas may be metal poor in which case the true H I covering fraction could be underestimated somewhat by our technique.

The relevant radiative transfer equation is:

$$\frac{I(v)}{I_0} = (1 - f_{\mathrm{cov}}(v)) + f_{\mathrm{cov}}(v) e^{-\tau(v)} \tag{2.1}$$

where I_0 is the continuum level and τ is the optical depth of each absorption line:

$$\tau(v) = \frac{\pi e^2}{m_e c} N(v) f_{lu} \lambda_{lu} = \frac{N(v)}{3.768 \times 10^{14}} f_{lu} \lambda_{lu} \tag{2.2}$$

where f_{lu} is an oscillator strength of each transition from level l to u (available from the NIST Atomic Spectra Databased), λ is the transition wavelength in Å, and $N(v)$ is the ion column density in $\mathrm{cm}^{-2}(\mathrm{km\ s}^{-1})^{-1}$. Equation 2.2 assumes none of the excited states are populated, i.e.,$N_u/N_l \approx 0$. Combining equations 2.1 and 2.2, the attenuation in the continuum normalized spectra at each velocity of an absorption line is then a function of two variables: $f_{\mathrm{cov}}(v)$ and column density of absorbers $N(v)$ with velocity v. If we have more than two absorption lines originating from the same ion, we can solve for the two variables.

This picture of a "picket fence" for the ISM, whereby each individual absorbing cloud forms a 'picket' for each velocity bin, is consistent with that determined to be an appropriate model for local analogs of high redshift star-forming galaxies. A good example follows observations of the Ca II, H and K absorption lines in clouds around hot OB stars that resolve into several radial velocity components (Adams, 1949). Moreover, Heckman, Borthakur, et al. (2011) found all four local Lyman Break Analogs with high residual intensities in the low-ionization absorption lines are a better match to the picket fence model compared to a uniform shell model.

Our DEIMOS spectra cover three absorption line transitions of Si II at 1260, 1304 and 1526 Å enabling us to solve for both $f_{\mathrm{cov}}(v)$ and $N(v)$. We bin the absorption lines into several velocity components with the width of each velocity bin chosen to be equal to the spectral resolution at that wavelength. We limit the usage of Si II λ1304 to only those systems whose velocity $v \gtrsim -200$ km s^{-1} in order to avoid contamination by O I λ1302. We use a brute force (grid search) technique to find the best fit parameters for each velocity bin. The likelihood of each pair of parameters is calculated from the least-square residual $\chi^2 = \sum(I_{\mathrm{obs}} - I_{N,f_c})^2$. We adopt a prior range for $f_{\mathrm{cov}}(v)$ of [0,1] while the prior for $N(v)$ is adjusted so the range at each velocity covers the whole posterior probability distribution. An exception is made is when the lines are optically thick, in which case only a lower limit on the column density can be obtained. Examples of posterior distributions for both optically thin and optically thick regimes are shown in Figure 2.3.

Method 2: Average Low-ionization Absorption Profiles

In our second method, we simplify the covering fraction model given in Section 2.4 by assuming that all clouds are optically thick $\tau \gg 1$. Equation 2.1 then becomes

$$f_{\mathrm{cov}}(v) = 1 - \frac{I(v)}{I_0} \tag{2.3}$$

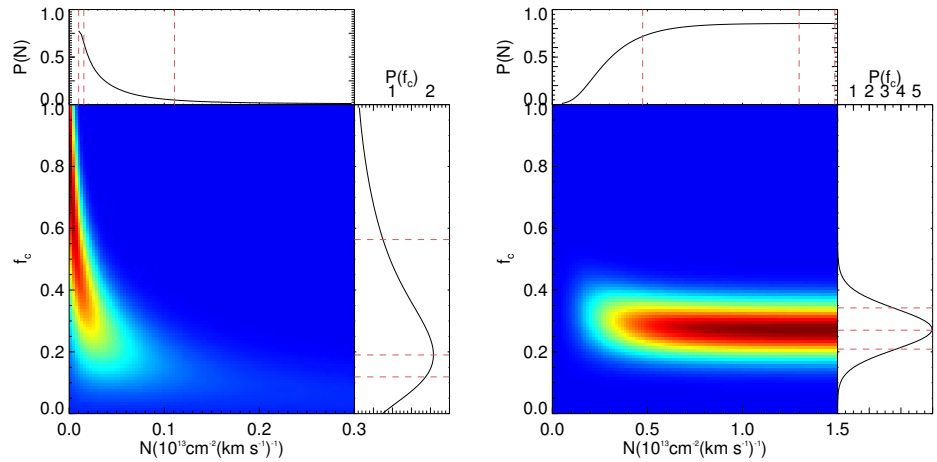

Figure 2.3: Posterior distributions of covering fractions ($f_{\rm cov}$) and column densities(N) in MACS 0940 at velocity $v = -270 \pm 50$ km s^{-1} (left) and at $v = -20 \pm 50$ km s^{-1} (right). The likelihood are calculated from three Si II transitions at 1260, 1304 and 1526 Å. The middle color plots represent 2-D posterior distribution with their marginalized 1-D posterior distributions of column densities and covering fractions projected on the top and on the right. The best values(median) and their 1σ uncertainties are shown in dashed lines. Most of the posterior distributions show optically thick absorption lines as shown in the right figure.

This enables us to include absorption lines from other species present in our spectra, including O Iλ1302 and C IIλ1334. We will see that the optically thick case is a reasonable assumption as results from Method 1 show that most of the Si II absorption lines are optically thick[3]. At each cloud velocity, we then calculate the covering fraction as an inverse-variance weighted mean of the line residuals.

Escape Fraction Constraints

The covering fractions estimated from low-ionization absorption profiles at each velocity resolution element are shown in Figure 2.4. In general we find that the absorbing gas is spatially inhomogeneous: the absorption line profiles show only partial covering of the stellar radiation with a maximum absorption depth of ∼50–100%. Formally this translates to geometric escape fractions of up to 50%. Only two of the seven lensed galaxies, A2219 and J1621, are consistent with a complete (100%) covering fraction at any velocity and hence a zero escape fraction. Without accounting for possible spatial variations in the covering fraction as a function

[3]Si II 1260, 1304, and 1526 are optically thick when N(Si II) $\gtrsim [0.02, 0.31, 0.18] \times 10^{13}$ cm^{-2}. O Iλ1302 and and C IIλ1334 are optically thick when N(OI)$\gtrsim 0.56 \times 10^{13}$cm^{-2} and N(C II) $\gtrsim 0.22 \times 10^{12}$cm^{-2} respectively.

of velocity, which will be examined in Section 2.4, we find a median value of covering fraction in our enlarged sample is $f_{\rm cov,obs} = 66\%$. If this represents the true H I covering fraction, corresponds to $f_{\rm esc,rel} = 34\%$. However, to translate these measures to the desired absolute escape fractions, it is necessary to account for dust extinction since these galaxies are not dust-free as indicated by the fact their UV slopes $\beta > -2$ (see Table 2.2). Accounting for reddening based on the dust-in-cloud model (see Vasei et al., 2016), the corresponding median $f_{\rm esc,abs}$ and weighted standard deviation of our sample becomes $19 \pm 6\%$.

At face value, such an absolute escape fraction (if typical of sources at higher redshift) would imply that early star-forming galaxies are readily capable of maintaining reionization(e.g. Robertson et al., 2015). Our median value derived from low-ionization absorption profiles is consistent though somewhat larger than that inferred via recombination rate studies of the IGM using QSO absorption lines at a similar redshift, which imply $f_{\rm esc} \sim 12\%$ (G. D. Becker et al., 2015). If $f_{\rm esc}$ is indeed 12% then we would expect an average $f_{\rm cov,obs} \sim 78\%$ for our sample, accounting for reddening. It is likely that our estimate is higher than the true value due to several violations of the assumed simple "picket fence" model of absorption as discussed in T. A. Jones, R. S. Ellis, et al. (2013) and Vasei et al. (2016). We now explore these possibilities quantitatively in order to estimate the degree by which our $f_{\rm cov,obs}$ measures may overestimate the true escape fraction.

Firstly, it is possible that the non-ionizing stellar radiation could be associated with a lower covering fraction of gas than that for the ionizing starlight. This would arise if, for example, the H I and low-energy ions preferentially cover younger stars. One way to estimate the potential importance of this effect is to estimate the fraction of detected UV continuum radiation (at ∼1500 Å) which arises from starlight whose photons are incapable of ionizing hydrogen. For a galaxy with a history of continuous star formation over 100 Myr, only around 10% of the UV radiation arises from spectral types later than B (Parravano et al., 2003) so this cannot be a significant effect. Furthermore the Si II transitions show no evidence of a higher covering fraction at shorter wavelengths. Such an effect would result in weaker absorption of Si II λ1526 Å compared to λ1304 Å in an optically thick gas, which is not observed in the majority of our combined sample where both lines are free from sky contamination. We therefore conclude that a differential covering fraction of ionizing stars introduces $< 10\%$ systematic error on our derived values.

A second uncertainty arises from the fact that scattered photons are capable of

"filling in" the absorption profiles, mimicking a reduced covering fraction. In the extreme case where line filling is induced from optically thin gas, the Si II 1304 and 1526 Å absorption profiles would be filled by 50% of the flux seen in the Si II* 1309 and 1533 lines, respectively. The observed fine structure emission line fluxes therefore allow us to directly estimate the magnitude of this effect. In Abell 2219 where these emission lines are strongest, we find that scattering could only reduce the inferred covering fractions by at most 13%. On average across the full sample, we find line filling can change our derived $f_{\rm cov,obs}$ by at most 4% (i.e.,to $f_{\rm cov,obs}$ = 70%) which is small compared to the quoted uncertainties. The actual effect of line filling must be even less than this limit as most cases indicate optically thick absorption.

In summary, accounting for the possibilities of wavelength-dependent covering fraction and line filling discussed above, we can revise the original covering fraction to at most $f_{\rm cov,obs} = 66 + 10 + 4 = 80\%$ implying $f_{\rm esc,abs} \sim 11\%$. However, we emphasize that these corrections represent the maximum conservative limits allowed by our data, and there is no conclusive evidence that the effects discussed actually reduce $f_{\rm esc}$ at all. Therefore, the difference between our covering fraction measurement and the $f_{\rm esc,abs} \sim 12\%$ implied by IGM studies, if real, can only be reconciled if these are the only important effects.

The most challenging and remaining uncertainty arises from the possibility that absorbing gas at different velocities may cover different spatial lines of sight (e.g. Rivera-Thorsen et al., 2015). To examine this possibility, we would need to examine the *s*patial variation of the absorption line profiles. Normally this would currently be impractical in any source at $z > 4$ but, by good fortune, this is possible in our most distant lensed source MS1358. In the next section, we show that it is likely that such spatial variations may be the largest cause of difference between our derived covering fractions and the true ionizing escape fraction.

Spatial Variations in the Low Ionization Gas

In the specific case of the highly-magnified target MS1358 at $z = 4.93$, we can examine the covering fraction of low ionization gas in discrete sub-components of the galaxy. When the region being examined spectroscopically is physically small, the angular distribution of clouds of different velocities should be less of a concern. In addition, if a particular region has a high star-formation rate, the possibility that its stellar continuum is affected by non-ionizing radiation should also be reduced. By

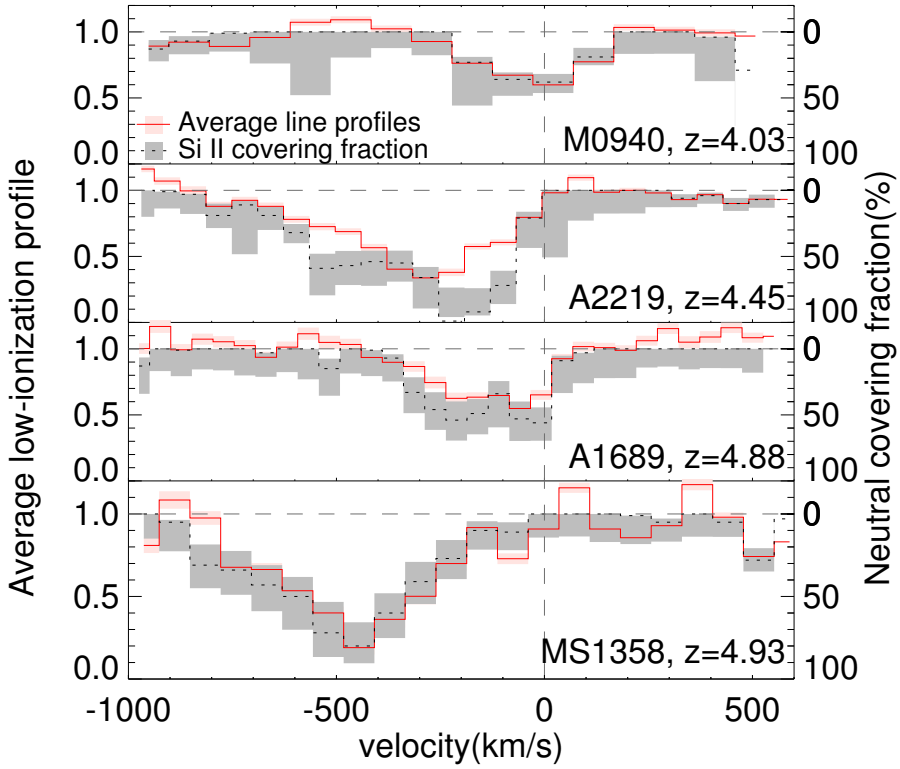

Figure 2.4: Average low-ionization absorption line profiles and covering fraction derived with Method 1 (black lines) and Method 2 (red lines) in Section 2.4. The shaded areas show 1σ uncertainties for each method. Each velocity bin corresponds to one spectral resolution. With no dust assumption, the covering fraction ranges from $\sim 50\% - 100\%$ at the central velocities.

examining the spectroscopic variation among the star-forming clumps in MS1358, we can therefore address several of the uncertainties in the inferred escape fraction from our methods discussed above.

Adopting the lensing magnification from our LTM model (Section 2.3) and physical parameters for MS1358 calculated in Swinbank, Webb, et al. (2009), its total stellar mass is $\sim 7 \pm 2 \times 10^8 M_\odot$ and its integrated star-formation rate(SFR) is $\sim 75 \pm 15 M_\odot$ yr^{-1}. The source itself is only a few kpc^2 in size. Since the individual clumps are conveniently linearly distributed along the arc, we have extracted spectra of each of 3 regions for specific clumps identified by Swinbank, Webb, et al. (2009). Due to the limited spatial resolution of our ground-based spectroscopy and the signal to noise of our data, we group 3 of the smaller clumps (C1-3) together in extracting the spectra. The resulting arrangement targeting 3 regions (C0, C1-3 and C4) is shown in Figure 2.5 alongside the [O II] image derived from the Swinbank, Webb, et al.

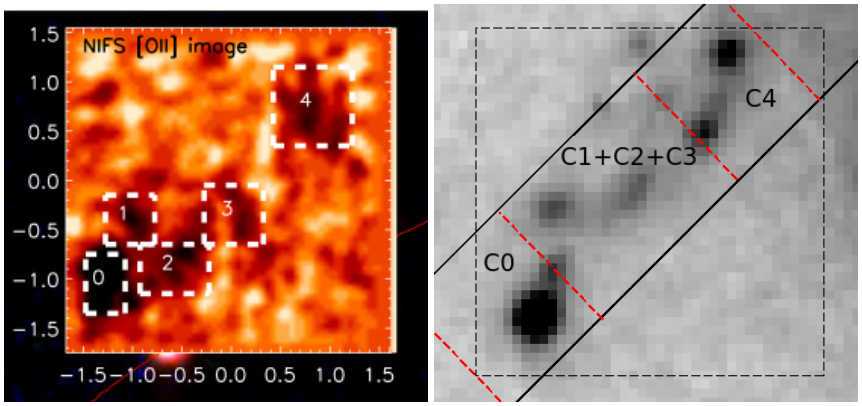

Figure 2.5: Resolved spectroscopy in the highly-magnified source MS1358 at $z = 4.93$. (Left) Distribution of [OII] emission from the integral field observations of Swinbank, Webb, et al. (2009). (Right) HST ACS I-band image overlaid with the 1 arcsec DEIMOS slit (solid black) and the Swinbank, Webb, et al. (2009) integral field image (dashed black). DEIMOS spectra were extracted for three regions containing clumps C0, C1+C2+C3 and C4 following the nomenclature of Swinbank, Webb, et al. (2009) (also marked in the left image).

(2009) integral field unit observations.

Importantly, we find there is a significant variation in the low-ionization absorption profiles across the source (Figure 2.6). The inferred covering fraction ranges from near-complete (100%) in C0 to only $\simeq$ 40% in C4. The variation within MS1358 is comparable to that across the integrated analysis of our full sample shown in Figure 2.4. As expected, it suggests that the escape fraction is governed by small-scale physics associated with star-forming regions. The variation between individual clumps indicates that covering fractions derived from integrated spectra may not be representative of the sum of all regions, and we can examine the degree of this effect. The SFR-weighted total of individual clumps is $f_{\rm cov} \sim 0.9$ whereas we find $f_{\rm cov} \sim 0.8$ based on the integrated light (Figure 2.7). It is possible that even higher covering fractions would be found at finer spatial resolution, although we note that Borthakur et al. (2014) found $f_{\rm esc,rel} \approx 1 - f_{\rm cov}$ in a compact source of similar size to the clumps studied here ($\sim$ few hundred pc). We conclude that spatial variation in the covering fraction therefore reduces the limit on $f_{\rm esc}$ by a factor of $\simeq$ 2 in this source. To help determine the extent to which this source is representative of the larger sample, in the following sections we examine trends in our sample as a function both local and integrated physical properties.

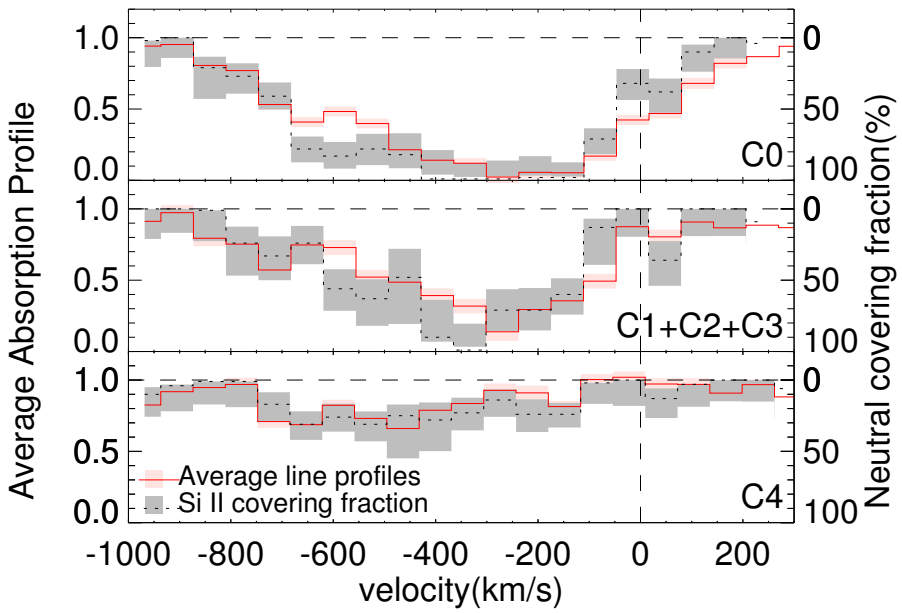

Figure 2.6: Average low-ionization absorption line profiles and covering fractions of *individual clumps* in the highly-magnified source MS1358. The variation in the maximum absorption depths seen within the source is comparable to that seen across the integrated analysis of our enlarged sample in Figure 2.4.

Variations with the Lyα Equivalent Width

Since both Lyα and LyC radiation are scattered by clouds of neutral hydrogen, the low column density region from which LyC can escape should also provide an escape route for Lyα photons. A suite of radiative transfer simulations in clumpy ISM models by Dijkstra, Gronke, et al. (2016) predicts a strong correlation between the escape fractions of Lyα and LyC. Indeed, observations indicate an anti-correlation between the low-ionization absorption depth and Lyα emission line equivalent width ($W_{Ly\alpha}$) (e.g. T. A. Jones, R. S. Ellis, et al., 2013; T. Jones et al., 2012; Erb, Steidel, et al., 2014; Shapley, Steidel, Pettini, et al., 2003). If this reflects the correlation between the Lyα and LyC escape fractions, not only does it imply that low-ionization absorption depths are a valuable proxy for the LyC escape fraction, but it also provides observational support for the clumpy ISM models.

We plot the relation between $W_{Ly\alpha}$ and the low-ionization absorption depth in Figure 2.7 combining both the integrated measures and those for the clumps identified in MS1358 (Figure 2.5). Most of the new data points support the anti-correlation found in T. A. Jones, R. S. Ellis, et al. (2013) as shown in dashed-line. However, two clumps within MS1358 (C0 and C1+C2+C3) appear as outliers although are consistent with

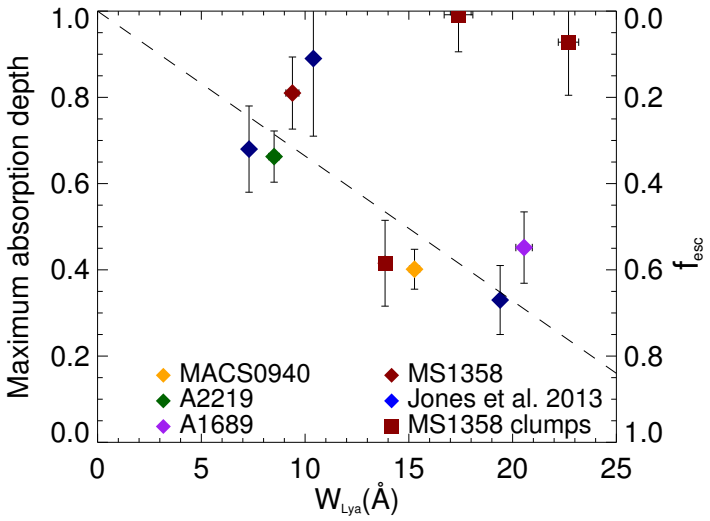

Figure 2.7: Correlation of the Lyα emission line equivalent width with the maximum low-ionization absorption depth, which supports the expected anti-correlation between the H I covering fraction and $f_{\rm esc}$. The dashed line show the earlier correlation published by T. A. Jones, R. S. Ellis, et al. (2013). With the exception of the intensely active clumps in MS1358, the new data supports the correlation. The scatter at high $W_{\rm Ly\alpha}$ is consistent with that seen at high $f_{\rm esc,Ly\alpha}$ found in simulations (Dijkstra, Gronke, et al., 2016).

the large scatter found in the Dijkstra, Gronke, et al. (2016) simulations. The origin of the large scatter in the simulations is the cloud-covering factor (the average number of clouds along each sight-line around the region). There is a higher Poisson probability that a line of sight would be free from obscuration when the cloud-covering factor is low. If we interpret Figure 2.7 accordingly to the simulation model, we can infer that the two clumps in MS1358 have a higher number of clouds given their increased star-formation rates.

Variations with Reddening

Reddy et al., 2016 find that galaxies in their $z \simeq 3$ sample with redder UV continua tend to have larger covering fractions of neutral hydrogen. They propose this correlation, modeled according to contributions from both photoelectric absorption and dust extinction, could form the basis of a new method to infer the escape fraction of high redshift galaxies, thereby bypassing the assumption of our method that the opacity of low ionization metal lines represent a reasonable proxy for that of neutral hydrogen. Although a correlation between the line of sight extinction, as inferred from the color excess $E(B-V)$, and $f_{\rm cov}$ seems reasonable, for the integrated spectra for our full sample, we nonetheless found no strong dependence between

our absolute covering fraction and $E(B-V)$. However, the three data points from the resolved data on MS1358 suggest that regions with higher covering fractions within the galaxy have redder UV continua and also stronger star formation rates as shown in Table 2.2. We will explore the variation of absorption line strength and star formation rates in the next section.

Variations with the Star Formation Rate

Our spatially-resolved observations of MS1358 presented in Section 2.4.5 suggested that the implied escape fraction of ionizing photons depends sensitively on the small-scale structure of its star-forming regions. Here we consider the origin of these variations in terms of both the local and integrated star-formation rates.

The star formation rate (SFR) may have two opposing effects on the escape fraction. First, the SFR is directly correlated with the column density of gas, including LyC-absorbing H I as well as molecular H_2 and dust (Krumholz et al., 2012; Bigiel, Leroy, et al., 2008; Bigiel, Bolatto, et al., 2010) However, strong feedback and radiation pressure in high SFR regions may clear the absorbing clouds resulting in a higher escape fraction. Understanding the balance between these two effects in early galaxies is critical in resolving the question of whether galaxies are capable of maintaining an ionized IGM.

According to our data shown in Figure 2.8, we see tentative evidence of a modest anti-correlation between SFR and the inferred escape fraction. Formally this trend is non-zero at the 2.8σ level, with a correlation coefficient of -0.53 ± 0.19 for the galaxy-integrated measurements. The trend is apparently stronger (though statistically consistent) for the individual clumps in MS1358 as compared to that seen for the integrated values across the galaxies in the sample. If real, this difference might be expected since the escape fraction inferred for an entire galaxy is necessarily a spectroscopic measure averaged over clumps with a range of both low and high SFRs. The observed anti-correlation can be interpreted in two independent ways. Firstly, regions of low SFR are associated with lower column densities of both interstellar and outflowing material (including H I and dust). Therefore, the probability of holes permitting leaking LyC photons is higher. Alternatively, regions with lower SFR may represent those whose rates were higher in the past and are now in decline. The associated delay time increases the chance that feedback effects can clear any covering clouds. Although both explanations may be valid, we consider the latter more likely because only a small column density of H I gas is needed to extinguish

any leaking LyC radiation. Regardless of the surface density of low ionization gas, it seems feedback will be the more effective factor in governing both a low SFR and a smaller H I covering fraction.

Most radiative transfer hydrodynamical simulations of the resolved ISM support our conjecture that a time delay will increase the escape fraction. Star formation may become increasingly 'burst-like' at high redshift(Stark, R. S. Ellis, A. Bunker, et al., 2009) and the escape fraction will then fluctuate on short time scales, out of phase with the SFR. Kimm et al. (2014) find a delay of $\simeq$10 Myr for a covering cloud to be dispersed by feedback from supernovae, while Ma, Kasen, et al. (2015) found a slightly shorter delay of $\simeq$3-10 Myr. A larger delay will increase the strength of the anti-correlation between SFR and escape fraction.

One cautionary remark is relevant in considering how to interpret Figure 2.8 given the SFR values are inferred from gravitationally-lensed galaxies using our lensing model (Section 2.3). While the intrinsic SFR can be reliably estimated by dividing the observed value by the appropriate magnification, the covering fraction is measured from the absorption line depths in the image plane. Hence, there is a potential bias such that the measured covering fraction will refer to that of the regions with the highest magnifications. It would only be possible to appropriately transfer the absorption line spectra into the source plane with resolved spectroscopy using an integral field unit spectrograph. Such a bias would, however, only be significant if there is a significant magnification gradient along an arc, e.g.. when it is close to a critical curve. In our sample, this issue is not important. Both A1689 and A2219 have low and uniform magnifications and the individual MS1358 clumps are sufficiently small as to have uniform magnifications within each region. When MS1358 is considered as an entirety, we excluded arc B (Figure 2.1) which has a high variation in magnification. MACS0940 is the only system where the variation in magnification is significant. However, its SFR is relatively low so such a bias is unlikely to affect its location on the relation in Figure 2.8.

Redshift Evolution and Reonization

It is clear that, despite the significant uncertainties, our moderate covering fractions imply absolute escape fractions that are higher than the averages observed via direct measures of the LyC radiation at lower redshift (e.g. Mostardi et al., 2015). On the other hand, if star formation is increasingly burst-like at high redshift where galaxies are physically more compact, it is quite likely from the arguments discussed above,

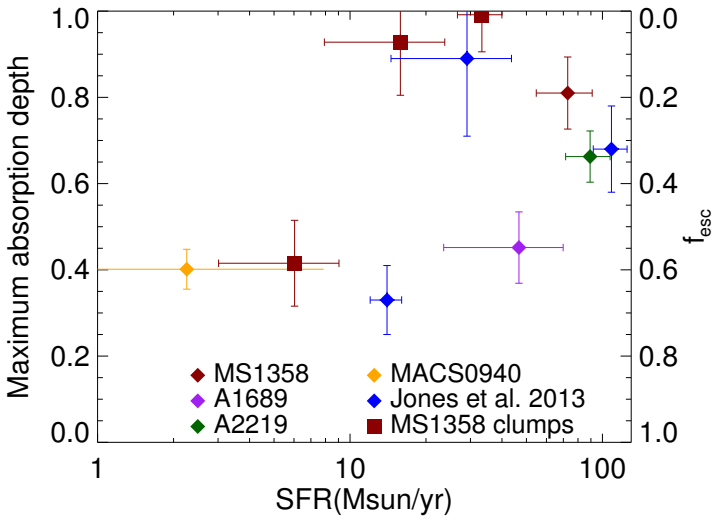

Figure 2.8: The correlation between the intrinsic star formation rate (after correction for the local or integrated magnification) and the maximum low-ionization absorption depth and covering fraction. Galaxies with high star formation rates tend to have higher covering fractions (and therefore lower value of $f_{\rm esc}$). The trend is apparently stronger in individual star-forming clumps.

that there is increase in $f_{\rm esc}$ with redshift. Theoretical studies present conflicting predictions on the question of evolution in $f_{\rm esc}$. Ferrara et al. (2013) present an analytical calculation that illustrates why the average escape fraction should decrease with time during the reionization era. With an assumed value for the efficiency of star-formation, they estimate $f_{\rm esc}$ decreased by a factor of 3 from $z = 10$ to $z = 6$. However, sample-limited radiative hydrodynamical simulations show no significant variation over the same redshift interval (Ma, Kasen, et al., 2015).

In compiling the data for Figure 2.9 we include observations at redshifts $z < 4$ from T. A. Jones, R. S. Ellis, et al. (2013) (not shown in Table 2.1), Reddy et al. (2016), and Patrício et al. (2016). Across this increased redshift range, the figure shows that we cannot yet detect any significant time dependence. This is because the scatter in covering fraction is dominant and apparently constant over time, in agreement with our deduction that the escape fraction depends largely on the small-scale structure of clouds, their physical association with star formation regions and the short timescale of activity, and is consistent with the behavior of simulated galaxies in Ma, Kasen, et al. (2015).

As shown in Figure 2.9, absorption line measurements of comparable quality to those presented here are still scarce at z = 2-3. Recent studies of the “cosmic horseshoe"

at $z = 2.4$ further highlights the challenge in comparing direct and indirect methods. The covering fraction was found to be $\sim$ 60% using methods equivalent to ours (Quider et al., 2009), yet imaging below the Lyman limit indicates $f_{\rm esc,rel} < 0.08$ (Vasei et al., 2016). This might be reconciled if $f_{\rm cov}$ overestimates $f_{\rm esc}$ by a factor of 5, or by low IGM transmission (~20% probability, Vasei et al. (2016), or a combination of both effects. At present the cause is not clear and it will be of great interest to examine low-ion covering fractions in more low redshift sources with direct $f_{\rm esc}$ measurements. This is becoming feasible following recent LyC detections in nearby galaxies using HST/COS (e.g. Izotov, Orlitová, et al. (2016) and Leitherer et al. (2016)). Reddy et al. (2016) present new measures of the covering fraction of both the low ionization gas and neutral hydrogen for stacked spectra derived from their Keck survey at $z \simeq 3$ providing a valuable complementary approach to that discussed here. Although we have highlighted the possible difficulties of inferring a mean covering fraction from a stack of many spectra with diverse kinematic absorption line profiles (such as the variety shown in Figure 2.4), Reddy et al. (2016) draw from a much larger sample and have questioned some of the assumptions inherent in the use of low ionization gas alone. However, within the observational uncertainties at this stage, both approaches give reasonably consistent results for the inferred escape fraction at $z > 3$.

Despite significant observational and theoretical effort, direct evidence for a rising escape fraction at high redshift is still lacking due to the scarcity of LyC detections. Our data set suggests a median $f_{\rm esc,abs} = 19 \pm 6\%$ at z = 4-5 if we interpret the covering fraction as $f_{\rm cov} = 1 - f_{\rm esc,rel}$. However, spatial variation may lower this value to $f_{\rm esc,abs} \approx 10\%$ if the resolved properties of MS1358 are typical of our sample, and other effects discussed in Section 4.4 may reduce the value slightly further (to $\sim$ 7%). These estimates are certainly compatible with independent determinations from the IGM, as well as an increase over the limits at $z \sim 2$. Furthermore the demographic trends in Figures 2.7 and 2.8 indicate that we may expect escape fractions to increase with redshift, commensurate with increasing $W_{{\rm Ly}\alpha}$ (e.g. Law, Steidel, Shapley, et al., 2012). We speculate that a key difference may be that star formation is more burst-like (e.g. higher specific SFR; Schreiber et al. (2015) at higher z. Intense short-duration starburst episodes may be capable of clearing sightlines through the local ISM resulting in increased ionizing escape during periods of lower activity, giving rise to the trend in Figure 2.8.

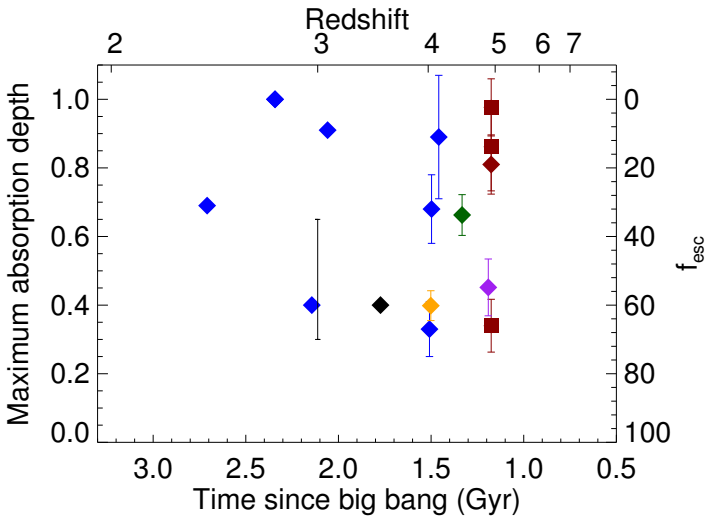

Figure 2.9: Time-dependence of the maximum low-ionization absorption depth with cosmic time for the present and earlier samples from T. A. Jones, R. S. Ellis, et al. (2013) and Reddy et al. (2016) spanning the redshift range $2.5 < z < 5$. No trend is apparent and the scatter supports our deduction that the escape fraction is governed by the small scale behavior of star-forming regions which fluctuate in their activities over short timescales. The color scheme follows that in Figure 2.8. The black error bar and the black data point show $f_{\rm cov,obs}$ derived from low-ionization absorption profiles, in a stacked spectrum of $z \sim 3$ galaxies from Reddy et al. (2016) and in a spectrum of a lensed $z = 3.5$ galaxy from Patrício et al. (2016) respectively.

2.5 Summary

We present intermediate resolution absorption line spectroscopy for four new gravitationally-lensed sources which, together with our earlier work (T. A. Jones, R. S. Ellis, et al., 2013), provides a total sample of seven star-forming galaxies in the redshift interval $4 < z < 5$. Our signal to noise in all cases is adequate to sample the depth of absorption in low ionization species which can act as a valuable indicator of the covering fraction of neutral hydrogen. We describe two methods to convert this covering fraction into estimates of the fraction of Lyman continuum photons that are capable of reaching the intergalactic medium and undertake a critical analysis of possible uncertainties in our results. Our enlarged sample strengthens the trends discovered in our earlier work. In particular we find:

- The amount of absorbing gas varies significantly from galaxy to galaxy across our sample. However, only two systems are consistent with a complete coverage of low ionization gas. The median observed covering fraction is 66%. Following a correction for reddening based on a dust-in-cloud model, we infer

a median absolute escape fraction of Lyman continuum photons of 19 ± 6%

- We discuss various uncertainties in this estimate including the possibility that the stellar radiation has different covering fractions depending on its ionizing capability as well as scattered photons which could fill in the absorption line profiles mimicking a reduced covering fraction. We demonstrate that, even in the extreme case, these issues would only lower the escape fraction to no less than 11%, a value in broad agreement with independent estimates based on recombination studies of the IGM probed by QSO absorption lines. Further study of low-ion absorption lines in nearby galaxies with recent LyC detections will help determine the degree to which covering fractions may overestimate $f_{\rm esc}$.

- For one of our lensed sources, MS1358, we demonstrate that the low ionization gas is clearly spatially inhomogeneous with covering fractions ranging from 40-100% across various star-forming clumps. Although our sample generally strengthens the correlation we found earlier whereby the escape fraction increases with the equivalent width of Lyα emission, the scatter increases when we consider trends in the individual star-forming clumps in MS1358, consistent with recent numerical simulations. As the escape fraction is anti-correlated with the local star formation rate, this suggests a time delay between a burst and the emergence of leaking Lyman continuum photons.

- Overall, our observations indicate significant variations in the escape fraction, both from galaxy-to-galaxy and spatially within an individual galaxy. This pattern is consistent with the ability of escape being governed by the small scale behavior of star-forming regions which fluctuate in their activities over short timescales. This conclusion supports the suggestion that the escape fraction may increase toward the reionization era when star formation becomes more energetic and burst-like in nature.

We acknowledge useful discussions with Phil Hopkins, Evan Kirby, Xiangcheng Ma, Brant Robertson, Brian Siana and Alice Shapley. TAJ acknowledges support from NASA through Hubble Fellowship grant HST-HF2-51359.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. RSE acknowledges support from the European Research Council through an Advanced Grant FP7/669253. DPS acknowledges support from the National Science

Foundation through the grant AST-1410155. AZ is supported by NASA through Hubble Fellowship grant #HST-HF2-51334.001-A awarded by STScI, which is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. Data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology. the University of California and NASA. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors acknowledge the very significant cultural role that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Chapter 3

A KECK SURVEY OF GRAVITATIONALLY-LENSED STAR-FORMING GALAXIES: HIGH SPATIAL RESOLUTION STUDIES OF KINEMATICS AND METALLICITY GRADIENTS

Leethochawalit, N. et al. (2016), ApJ 820, 84, 84

Abstract

We discuss spatially resolved emission line spectroscopy secured for a total sample of 15 gravitationally lensed star-forming galaxies at a mean redshift of $z \simeq 2$ based on Keck laser-assisted adaptive optics observations undertaken with the recently-improved OSIRIS integral field unit (IFU) spectrograph. By exploiting gravitationally lensed sources drawn primarily from the CASSOWARY survey, we sample these sub-L* galaxies with source-plane resolutions of a few hundred parsecs ensuring well-sampled 2-D velocity data and resolved variations in the gas-phase metallicity. Such high spatial resolution data offers a critical check on the structural properties of larger samples derived with coarser sampling using multiple-IFU instruments. We demonstrate how kinematic complexities essential to understanding the maturity of an early star-forming galaxy can often only be revealed with better sampled data. Although we include four sources from our earlier work, the present study provides a more representative sample unbiased with respect to emission line strength. Contrary to earlier suggestions, our data indicates a more diverse range of kinematic and metal gradient behavior inconsistent with a simple picture of well-ordered rotation developing concurrently with established steep metal gradients in all but merging systems. Comparing our observations with the predictions of hydrodynamical simulations suggests that gas and metals have been mixed by outflows or other strong feedback processes, flattening the metal gradients in early star-forming galaxies.

3.1 Introduction

Statistical surveys of star forming galaxies at the peak era of activity, corresponding to a redshift interval $z \simeq$1.5–3, have been increasingly complemented over the past few years by spatially resolved spectroscopic observations. Early work focused on characterizing global trends such as the evolution of star formation rate density with redshift (Madau et al., 2014), and evolution of the “main sequence" of star formation

rate (SFR) as a function of galaxy stellar mass (Noeske et al., 2007; Y.-j. Peng et al., 2010). Spatially resolved data from integral field unit (IFU) spectrographs has provided complementary information on the ionized gas kinematics (Förster Schreiber, Genzel, Lehnert, et al., 2006; Stark, Swinbank, et al., 2008; Law, Steidel, Erb, et al., 2009; Förster Schreiber, Genzel, Bouché, et al., 2009; T. Jones, R. Ellis, et al., 2010; Wisnioski, Glazebrook, et al., 2011; Swinbank, Sobral, et al., 2012; Wisnioski, Förster Schreiber, et al., 2015), the size and spatial distribution of giant star forming regions (Genzel, Burkert, et al., 2008; T. A. Jones, Swinbank, et al., 2010; Livermore, T. Jones, et al., 2012), and radial metal abundance gradients (Cresci et al., 2010; T. A. Jones, Swinbank, et al., 2010; Yuan, Kewley, Swinbank, et al., 2011; Swinbank, Sobral, et al., 2012; T. A. Jones, R. S. Ellis, et al., 2013; Stott et al., 2014; Troncoso et al., 2014; T. Jones, Wang, et al., 2015). 2-D kinematic data and the properties of clumpy star forming regions provide valuable insight into the emergence of primitive disks, the role of instabilities, and the gradual assembly of central bulges possibly formed from inwardly migrating clumps (Genzel, S. Newman, et al., 2011). The properties of these clumps provides important evidence of gravitational instabilities in the primitive disks, whereas correlations between their ages and radial positions insight into the morphological evolution of these early systems (Förster Schreiber, Shapley, et al., 2011).

Metallicity gradients also provide an important opportunity to study the amount of feedback, i.e., the energy returned to the surrounding medium by star formation and nuclear activity. Energy input can arise from various mechanisms including supernovae, radiation pressure, and cosmic rays. These effects combine to drive large-scale outflows of gas and metals. Outflows are ubiquitously observed from systems with high star formation rate densities (Heckman, Sembach, et al., 2001), including virtually all star forming galaxies at high redshifts (e.g. Shapley, Steidel, Pettini, et al., 2003). Cosmological simulations confirm that outflow mass loss rates must be comparable and often larger than star formation rates in order to explain measured stellar mass functions and the mass-metallicity relation (e.g. Vogelsberger et al., 2014; Crain et al., 2015). The rates of mass and metal loss are degenerate with rates of subsequent accretion and hence are poorly known. However, this “galactic fountain" cycle may be constrained via its imprint on metallicity gradients. To first order, high rates of outflow and subsequent accretion will redistribute heavy elements, resulting in a flatter gradient. Furthermore, the redistribution of interstellar gas will affect the radial profile of star formation and hence future metal production and feedback. Evidence for these effects has been recently observed in the form

of flattened metal gradients in the extended disks of nearby galaxies (e.g. Bresolin, 2011). Several groups have now explored how various forms of feedback affect metal gradients in simulations (e.g. Yang et al., 2012; Gibson et al., 2013; Anglés-Alcázar et al., 2014). Despite a range of methods, there is a clear consensus that stronger feedback (that is a higher rate of energy injection from stellar winds, supernovae, and other sources) results in flatter gradients due to mixing of gas and metals over larger physical scales. Various prescriptions are able to match observations of local galaxies but predict different behavior at early times. We therefore seek to constrain the degree of gas cycling via feedback, by directly measuring gradients at high redshifts.

Initial IFU surveys targeted modest numbers of galaxies beyond $z \simeq 2$ (Förster Schreiber, Genzel, Lehnert, et al., 2006; Law, Steidel, Erb, et al., 2009; T. Jones, R. Ellis, et al., 2010), reflecting the challenges of securing adequate signal to noise per spatial resolution element. The advent of multi-IFU spectrographs (e.g., KMOS on the ESO VLT; Sharples et al. 2013) has recently led to a significant improvement in survey capability. For example, Wisnioski, Förster Schreiber, et al. (2015) report resolved kinematic data for over 100 galaxies within the redshift range $2 < z < 2.7$ using this impressive instrument. However, the multi-IFU approach comes with a major limitation in terms of angular resolution. Typical $0''.6$ seeing conditions correspond to a physical scale of 5 kpc at $z \simeq 2$, whereas the half-light radius of a typical L^* galaxy at this redshift is only 2 kpc. Clearly only the largest and most massive systems can be adequately probed with seeing-limited data. In fact, even with adaptive optics, the physical resolution is only $\simeq$1 kpc (Genzel, Tacconi, et al., 2006; Law, Steidel, Erb, et al., 2009). T. Jones, R. Ellis, et al. (2010) and T. A. Jones, R. S. Ellis, et al. (2013) illustrate the difficulties of correctly interpreting velocity fields and abundance gradients from data with so few resolution elements across each galaxy. As an example, early IFU surveys suggested that compact galaxies at $z \simeq 2$ were characterized by dispersion-dominated kinematics with little or no rotation (Förster Schreiber, Genzel, Bouché, et al., 2009). Deeper observations with adaptive optics showed that most of these sources in fact harbor rotating thick disks, with circular velocities commensurate with their lower masses compared to rotation-dominated systems (S. F. Newman, Genzel, Förster Schreiber, et al., 2013), confirming earlier results from lensed galaxies with superior resolution (T. Jones, R. Ellis, et al., 2010).

Studying gravitationally lensed systems with adaptive optics represents a highly

valuable route to addressing the physics of galaxy formation, particularly for the less massive and more abundant systems at $z > 2$. Lensing magnification enables higher spatial resolution and better sampling than is otherwise possible. In a pioneering study, Stark, Swinbank, et al. (2008) illustrated the potential of the lensing approach by securing resolved kinematic data for MACSJ2135-0102, a z=3.075 star-forming galaxy magnified in angular size by a factor of $\times$8 along its major axis by a foreground galaxy and galaxy cluster. Using the Keck OSIRIS IFU spectrograph with adaptive optics, the source-plane velocity field was sampled with a resolution of 120 pc leading to 20 independent points on its rotation curve. Subsequently, T. Jones, R. Ellis, et al., 2010 studied a sample of 6 lensed galaxies, demonstrating that a high fraction of such $\lesssim L^*$ systems display well-ordered velocity fields, in contradiction to the claims of earlier less well-sampled studies of unlensed galaxies of similar luminosities (Law, Steidel, Erb, et al., 2009). Likewise, T. A. Jones, Swinbank, et al. (2010) derived a radial metal abundance gradient for a lensed $z = 2.00$ galaxy utilizing the spatial variation of [NII]/Hα and [OIII]/Hα on 500 pc scales for a system with a half-light radius of 2.9 kpc. The necessity of mapping weaker lines such as [NII] makes the metal gradient aspect particularly challenging, but four systems were studied in T. A. Jones, R. S. Ellis, et al. (2013) showing a diverse range of gradients, possibly dependent upon the kinematic properties (see also Cresci et al. (2010)). T. A. Jones, R. S. Ellis, et al. (2013) introduced a simple model which suggests that metal gradients and their evolution should provide insight into the radial variation in the mass loading factor governing the amount of outflowing gas. More recently, a KMOS study of a larger sample of 21 galaxies at $z \simeq 1$ suggests the metal gradient tends to be less steep in more intensely star-forming systems (Stott et al., 2014). In view of the diverse results, clearly larger samples are required, particularly at high redshift where only lensed galaxies provide the necessary physical resolution.

This paper presents the results of a significantly enlarged IFU survey of lensed galaxies. Two practical developments have motivated us to extend the original sample presented in T. A. Jones, R. S. Ellis, et al. (2013). Firstly, the number of lensed targets with known spectroscopic redshifts has increased following the CASSOWARY survey of lensed star-forming galaxies located primarily in the Sloan Digital Sky Survey (Stark, M. Auger, et al., 2013). The second is a substantial improvement in the performance of the OSIRIS IFU spectrograph following its transfer to Keck I. This exploits the more powerful center launch laser which offers a brighter and more compact LGS beacon, ensuring a much improved Strehl. The installation of a new grating also improved the instrumental throughput by a factor

of ×1.5–2.5 depending on wavelength.

The aim of the present survey is to exploit recent improvements at the Keck Observatory to extend the original campaign, obtaining IFU spectroscopy of a larger and more representative sample of lensed galaxies. The improved throughput has enabled us to discard the selection criterion adopted by T. A. Jones, R. S. Ellis, et al., 2013 whereby targets were initially pre-screened with a long-slit spectrograph (NIRSPEC) to ensure adequately strong emission lines, possibly biasing the sample to unusually active sources. Our goals are two-fold. First, we aim to measure the radial metallicity gradient on sub-kpc scales for a representative sample of $z \simeq 2$ galaxies and to examine further the origin of the diverse results obtained by various groups. T. A. Jones, R. S. Ellis, et al. (2013) and Yuan, Kewley, and Rich (2013) argued that much of the discrepancy in metal gradient measurements, e.g. Queyrel et al. (2012) might arise from poorly-sampled data, whereas Stott et al. (2014) suggested correlations with the specific star formation may be the cause. Second, we aim to characterize the kinematics and star formation on sub-kpc scales for a larger sample. Our well-sampled kinematics and star formation morphologies allows us to evaluate the utility and conclusions drawn from complementary larger surveys being undertaken, e.g. with KMOS (Stott et al., 2014; Wisnioski, Förster Schreiber, et al., 2015), at coarser ∼5 kpc resolutions.

A plan of the paper follows. In §2 we discuss the sample drawn primarily from the CASSOWARY survey and the relevant selection criteria. We present the OSIRIS spectroscopic observations and their initial reduction. §3 discusses the reductions of the resolved spectroscopic data into the source plane utilizing the available mass models for the foreground lenses. We discuss the kinematic properties in §4 and the metal gradients in §5. §6 discusses the overall results and we summarize our conclusions in §7.

3.2 Data

Sample and Observations

During several observing runs in 2013–2014, we observed 11 gravitationally-lensed sources at a mean redshift $z \simeq 2$ using the near-infrared integral field spectrograph OSIRIS (Larkin et al., 2006) on the Keck 1 telescopes with the laser guide star adaptive optics (LGSAO) system (Wizinowich et al., 2006). Prior to 2013, we observed four similar sources using OSIRIS with the Keck II AO system and presented the results in T. A. Jones, R. S. Ellis, et al. (2013). Because of the lower system throughput

at that time, the early four sources were pre-screened with the long-slit spectrograph NIRSPEC to have suitably bright [N II] and Hα emission lines so as to ensure a reasonable signal-to-noise with OSIRIS in practical integration times. After the transfer of OSIRIS to the Keck II telescope and the installation of a new grating in 2012 December, the AO Strehl was improved and the OSIRIS throughput increased by a factor of $\simeq \times 2$ (Mieda et al., 2014). Accordingly, to avoid any bias in selecting targets, we abandoned the earlier spectroscopic pre-screening method. The current sample therefore comprises 15 sources in total. Ten new sources were selected from the CASSOWARY catalog of star-forming lensed galaxies as presented in Stark, M. Auger, et al. (2013) based on their availability during the scheduled observation period, their rest-frame UV spectroscopic redshifts in the range $z = 1.5-2.5$, and the presence of suitably bright proximate tip-tilt guide stars. We selected an additional cluster-lensed galaxy (Abell 773) from Belli et al. (2013) with similar criteria as the CASSOWARY sample. The eleven new sources extend the UV luminosity and SFR range in T. A. Jones, R. S. Ellis, et al. (2013) with UV absolute magnitudes in the range $M_{UV} \sim -22$ to -18 and SFR in the range 1 to 80 $M_{\odot}yr^{-1}$ (uncorrected for dust). Most of the sources are in the sub-L^* regime, while a few extend slightly beyond the L^* magnitude limit ($M_{UV} < -21$). The total sample is presented in Table 3.1.

Using reddening-corrected star formation rates (SFRs) derives from both Hα and ultraviolet continuum measures (using the methods described below), we can compare the properties of our present sample with those in previous studies. Although there is a wide range overall (from $\simeq$5 to 250 $M_{\odot}$ yr^{-1}), the bulk of our sample have values in the range 10 to 100, with a median of $\simeq 40 M_{\odot}$ yr^{-1}. This is comparable to the rates observed in the unlensed surveys(Law, Steidel, Erb, et al., 2009; Förster Schreiber, Genzel, Bouché, et al., 2009; Wisnioski, Glazebrook, et al., 2011; Wisnioski, Förster Schreiber, et al., 2015). Near-infrared photometry, essential for deriving accurate stellar masses, are available for about half the objects in our sample (and hence are not listed in Table 3.1). However, they lie in the range log M_* $\simeq$9.0-9.6, which is significantly less than for those in the unlensed surveys which generally probe systems from 10^{10} to 10^{11} $M_{\odot}$. Our survey therefore samples star-forming galaxies somewhat above the main sequence as defined by Behroozi et al. (2013).

We closely followed the observing technique described in T. A. Jones, Swinbank, et al. (2010). Observations were undertaken in one or more near-infrared passbands (J,H, or K) to secure spatially-resolved data on the Hα and [N II] emission lines.

Table 3.1: Observation Log

ID	z	Coordinates RA DEC	Dates (MM/YY)	Filter	Lines	t_{exp} (ks)	FWHM (PSF)	FWHM (Source Plane)	μ	$A_{H\alpha}$ (mag)	SFR ($M_\odot yr^{-1}$)
This paper											
cswa11	1.41	08:00:12 +08:12:07	3/13,2/14	Hn2	Hα, N[II]	16.2	0″.21	1.5×1.9 kpc	1.9	–	99 ± 90^b
cswa15	2.16	10:09:01 +19:37:23	2/14	Kn2	Hα, N[II]	9	0″.53[a]	2.2×4.4 kpc	7.6	0.52 ± 0.10	42 ± 5
			2/14	Hbb	O[III], Hβ	3.6	–	–			
cswa19	2.03	09:00:03 +22:34:08	3/13,2/14	Kn1	Hα, N[II]	14.4	0″.14	0.4×1.2 kpc	4.3	0.46 ± 0.17	59 ± 11
			3/13	Hn1	O[III], Hβ	3.6	–	–			
cswa20	1.43	14:41:49 +14:41:22	02/14	Hn2	Hα, N[II]	3.6	0″.09	0.1×0.3 kpc	14	0.27 ± 0.34	6 ± 3
			2/14	Jn1	O[III], Hβ	2.7	0″.25	0.5×1.1 kpc			
cswa28	2.09	13:43:33 +41:55:13	2/14	Kn1	Hα, N[II]	3.6	0″.23	0.3×1.8 kpc	9.3	–	12 ± 11
cswa31	1.49	09:21:25 +18:10:11	12/14	Hn3	Hα, N[II]	10.8	0″.22	0.6×2.1 kpc	3.3	–	36 ± 33
cswa128	2.22	19:58:35 +59:50:53	9/13	Kn2	Hα, N[II]	5.4	0″.15	0.4×0.7 kpc	10	$1.96^{+0.23}_{-0.77}$	250^{+71}_{142}
			9/13	Hbb	O[III], Hβ	5.4	–	–			
cswa139	2.54	08:07:32 +44:10:51	2/14,12/14	Kc5	Hα, N[II]	10.8	0″.18	0.7×1.2 kpc	9.7	$0.38^{+0.36}_{-0.52}$	33^{+16}_{-24}
			12/14	Hbb	O[III], Hβ	5.4	0″.07	0.2×0.4 kpc			
			2/14	Jn3	O[II]	1.8	–	–			
cswa159	2.30	22:22:09 +27:45:25	9/13,12/14	Kc3	Hα, N[II]	7.2	0″.17	1.2×2.7 kpc	4.6	0.44 ± 0.95	53 ± 47
			12/14	Hn3	O[III], Hβ	3.6	–	–			
cswa165	2.13	01:05:20 +01:44:58	9/13,12/14	Kn2	Hα, N[II]	7.2	0″.16	0.1×0.2 kpc	42	$1.03^{+0.2}_{-0.58}$	7^{+2}_{-3}
			9/13,12/14	Hbb	O[III], Hβ	5.4	0″.20	0.1×0.2 kpc			
a773	2.30	09:17:57 +51:43:31	3/13, 2/14,12/14	Kc3	Hα, N[II]	12.6	0″.39	0.1×1.2 kpc	20.3	2.5 ± 0.25	30 ± 7
			3/13	Hn3	O[III], Hβ	7.2	0″.41	0.1×1.2 kpc			
Objects published in T. A. Jones, R. S. Ellis, et al. (2013)											
J0744	2.21	07:44:48 +39:27:26	1/08	Kn2	Hα, N[II]	9	0″.11	0.3×0.8 kpc	16	0.53 ± 1.2	$5.4^{+4.9}_{-1.8}$
			2/11	Hbb	O[III], Hβ	3.6	0″.08	0.3×0.7 kpc			
J1038	2.20	10:38:42 +48:49:19	2/11,3/11	Kn2	Hα, N[II]	9	0″.14	0.4×1.6 kpc	8.4	0.67 ± 1.30	38^{+37}_{-15}
			2/11	Hbb	O[III], Hβ	3.6	0″.14	0.3×1.7 kpc			
J1148	2.38	11:48:33 +19:29:59	2/11	Kn2	Hα, N[II]	9	0″.11	0.6×0.9 kpc	10.3	2.94 ± 1.05	210 ± 167
			2/11,3/11	Hbb	O[III], Hβ	3.6	0″.08	0.6×0.9 kpc			
J1206	2.00	12:06:02 +51:42:30	5/10	Kn2	Hα, N[II]	9	0″.18	0.5×3.0 kpc	13.1	1.22 ± 0.47	68^{+44}_{-24}
			5/10	Hbb	O[III], Hβ	3.6	0″.33	0.6×3.5 kpc			

The 2-D magnification μ is defined as the ratio between image plane flux and source plane flux. UV magnitudes and Hα fluxes are of the source planes. [a] tip/tilt star is a galaxy. [b] For galaxies with no observations in the Hβ band, their star formation rates are obtained with the weighted mean extinction $A_{H\alpha} = 0.8 \pm 1.0$ in the sample.

Where practical, we continued to observe Hβ, and the [O III] emission lines within a shorter wavelength band. The average spectral resolution is $R \simeq 3600$ which corresponds to $\simeq 3-6$ Å across the J, H, K bands. We used the 100 mas pixel scale which gives a field of view of at least 1.6×6.4 arcsec. We took short exposures of each tip-tilt star to center the position before moving to the lensed galaxy. Each exposure comprised a number of 15-minute sub-exposures with a ABAB dither pattern ABAB of increment $\sim 2-3$ arcsec ensuring that the target was present in all frames. Exposure times varied from $1-4$ hours. The seeing during the observations varied between $0.4''$ to $1.5''$. The median AO-corrected seeing was $0.17''$.

Data Reduction

We used the latest OSIRIS Data Reduction Pipeline (Larkin et al., 2006)[1] to perform dark subtraction and cosmic ray rejection, followed by a direct or scaled sky subtraction prior to spectral extraction, wavelength calibration, and telluric correction using faint standard stars. Adjacent exposures were used as sky reference frames. The data cubes from each exposure were finally combined using a σ-clipped mean.

Emission Line Fitting

We fit a Gaussian to the various emission lines to determine the line flux, velocity, and velocity dispersion. In all cases, Hα is the most prominent line and a key indicator of the velocity field and star-formation rate distribution. A four-parameter Gaussian curve is fit to the Hα line for each spatial pixel using a weighted χ^2 minimization procedure. In each data cube, we select a region devoid of emission lines and calculate the weight for each wavelength from a variance over this region, $w(\lambda) = V^{-1}(\lambda)$, for use in the Gaussian fitting. We require that each Hα detection must be above 5σ and, where necessary, we spatially smooth the data cube with a Gaussian kernel of FWHM 3 pixels to increase the signal-to-noise. Given a Hα detection, we fit a Gaussian of identical width and velocity to the [N II] emission. Where observed, we fit the [O III] and Hβ emission lines independently from the Hα and [N II] emission lines but with a similar procedure. The velocities and line widths obtained from the [O III] emission lines are consistent with those obtained from Hα to within the 1σ uncertainties. The intrinsic velocity dispersion is calculated by subtracting the instrumental resolution measured from OH sky lines $\sigma_{inst} \simeq 50$ km s^{-1} in quadrature from the best-fit line width.

[1] Data Reduction Pipeline Version 3.2

Flux Calibration, Extinction and Star Formation Rate

For each galaxy observed in the Hα observation band, we use the tip/tilt reference star to calibrate the absolute flux. We fit a PSF to the image of an integrated flux over wavelength of the star and obtain a 1D star spectrum from the derived PSF. The flux calibration is then calculated at Hα wavelength of the galaxy. The only exception is CSWA15 whose tip/tilt reference star is an extended galaxy and its flux calibration is measured from a UKIRT infrared standard star FS26 observed in the same night. We crossed checked all the Hα flux calibrations with the flux calibrations derived from standard stars (FS9, FS11, FS19, or FS26) taken at the end of the same nights. Flux calibration derived from these standard stars generally agree to within 25%.

Ideally we would do the same process to the Hβ observation band observations. However, there are only three galaxies with tip/tilt reference stars taken in the Hβ bands (CSWA139, CSWA159 and Abell773). We thus use the flux calibration from the infrared standard stars to calibrate the Hβ fluxes when the information from tip/tilt stars are not available. The flux calibrations derived from the three tip/tilt reference stars are in agreement with those from standard stars within 15%.

We follow T. A. Jones, R. S. Ellis, et al. (2013) in the calculation of Hα extinction and star formation rate. In short, we used the Balmer line ratios Hα/Hβ with Calzetti et al. (2000) reddening curve to calculate dust extinction $E(B-V)$. We then calculate the extinction of Hα, $A_{H\alpha} = 3.33E(B-V)$. The star formation rate is computed from total Hα flux corrected for extinction and lensing magnification with the Hα SFR relation from Kennicutt (1998). For the three galaxies with no observations in Hβ bands (CSWA11,CSWA28, and CSWA31), we use the weighted mean and the standard deviation of $A_{H\alpha}$ of the other 12 galaxies in our sample, $\overline{A_{H\alpha}} = 0.8 \pm 1.0$, to calculate the star formation rates. The resulting $A_{H\alpha}$ and SFR are listed in Table 3.1.

3.3 Source Plane Reconstruction

By observing gravitationally-lensed systems we can secure much higher spatial resolution and sampling for our targets than would otherwise be the case. Nonetheless, this gain in resolution is only possible by using accurate mass models for the lens system via which our observations in the image plane can be transferred into the (unlensed) source plane. The key to developing appropriate mass models is the correct identification of multiply-imaged systems ideally with spectroscopic redshifts. Given the variety of lenses, from SDSS galaxies to Abell clusters, surveyed in this

study, it has not been possible to adopt a uniform approach to constructing these mass models across our sample.

When Hubble Space Telescope(HST) images or Gemini Science Archive (GSA) images were not available, we took photometric BRI images in good seeing with the Echellette Spectrograph and Imager (ESI) on the Keck II telescopes. In some cases, e.g. CSWA159, mass models were already available in the literature (Dahle et al., 2013). Depending on the circumstances, three distinct methods were used in constructing new mass models as detailed below. The methods and imaging data employed to develop the mass models are summarized in Table 3.2.

We used the Light-Traces-Mass (LTM) method (Zitrin et al., 2015b) to develop mass models for CSWA11, CSWA15, CSWA19, CSWA28, CSWA31, CSWA139, and CSWA159. Generally these lenses are dominated by a single galaxy with a few associated companions as they represent lenses with angular Einstein diameters in excess of the size of a SDSS fiber (> 3 arc sec). The method assumes that the lensing mass distribution of lensing galaxies is described by an elliptical power-law, with a gaussian core for the brightest central galaxies, scaled according to their luminosities. The member galaxies are identified spectroscopically and/or via a color cut. To represent the dark matter, the mass distribution of the galaxies is smoothed by a 2D Gaussian kernel. Multiple images are then identified from their similar colors and/or spectroscopic redshifts, if available, in conjunction with a preliminary guess for the mass model. In many cases, spectroscopic redshifts are available in the literature(Stark, M. Auger, et al., 2013; Dahle et al., 2013; Brewer et al., 2011) and from our OSIRIS observations. In a few cases where redshifts are not available for identified multiple images, we leave the redshift as a free parameter. In constructing the final best-fit mass model, we employ a Monte-Carlo Markov Chain (MCMC) to minimize the image position χ^2, i.e.,the total distance between the predicted and actual position of the multiple images. The model consists of six basic parameters: the exponent of the power law, the smoothing Gaussian width, the relative weight between the galaxy component and the dark matter component, the overall normalization of the mass distribution, and the two external shear parameters. The mass-to-light ratio is fixed for all galaxies except that for the brightest central galaxies that we left as a free parameter in each model.

Some mass models for CASSOWARY sources, CSWA20, CSWA128 and CSWA165 were already available following modeling developed by Matt Auger (MA) as part of that survey (Stark, M. Auger, et al., 2013). In these cases, *gri* imaging from

SDSS was used to fit each lensing component with a singular isothermal ellipsoid following the procedure discussed in M. W. Auger et al. (2011). The same mass model is fitted to each of the filters simultaneously, although the amplitudes of the surface brightness profiles can vary.

To determine how uncertainties in the lensing models affect our analyses of both the kinematic state of each galaxy and its chemical gradient, we empirically analyzed a marginal case of the lensing models. Since CSWA28 has a relatively high magnification factor and contains only one set of multiple images in constraining the lensing model i.e.,the magnification is the least well-constrained, we selected CSWA28 as the appropriate case. We constructed 20 different mass models following the parameters randomly drawn from the MCMC and performed the same kinematic properties analysis (simple disk model fitting) as in Section 3.4 below. We found that in most of the cases, different lens models primarily affect the position of the source relative to the image but the morphology and size of the galaxy in the source plane is largely the same. From the kinematic model fitting in each source plane, the derived galaxy inclinations, position angles, and centers are in agreement given the uncertainties we quote for the best model. The resulting uncertainty in radius at each pixel is $\lesssim$ 10%. In a few rare cases, the result from kinematic model fittings yields a radius that is 30% higher. For CSWA28, if we have propagated the expected maximum uncertainty of 10% in radii into the metal gradient calculation, the final uncertainty in the derived $N2_{PP04}$ metal gradient would have increased by only 2% from ±27% to ±29%. Since CSWA28 is one of the least certain lens models, we expect that the uncertainties due to lensing models are not the dominance source of uncertainty and did not propagate this uncertainty into the subsequent kinetic modeling.

Finally, for the rich cluster, Abell 773, the mass model is available from a detailed study conducted by Johan Richard (JR) (Richard, T. Jones, et al., 2011; Richard, Smith, et al., 2010; Limousin et al., 2012; Livermore, T. Jones, et al., 2012). In this case, a parametric mass model of the central region of the cluster was developed using the LENSTOOL package (Jullo et al., 2007). Briefly, we assumed the cluster mass distribution follows a double Pseudo-Isothermal Elliptical profile (Elíasdóttir et al., 2007) and added one or more central cluster members as smaller scale perturbations to the mass distribution. As above, optimization is conducted by minimizing the predicted and observed positions of the many multiple images in these well-studied clusters. Table 1 summarizes the overall 2-D magnification μ

Table 3.2: Mass Models

ID	Method	Image	Photometric bands
cswa11	LTM	Gemini	*gri*
cswa15	LTM	Keck ESI	BRI
cswa19	LTM	HST	GVIH
cswa20	MA	SDSS	*gri*
cswa28	LTM	HST	GVI
cswa31	LTM	Gemini	*gri*
cswa128	MA	SDSS	*gri*
cswa139	LTM	Keck ESI	VRI
cswa159	LTM	SDSS	*gri*
cswa165	MA	SDSS	*gri*
a773	JR	HST	YJH

appropriate for each source although, in the following analysis, full account is taken of the spatial dependence in transforming our data to the source plane.

3.4 Kinematic Properties

Methods and Motivation

The improved spatial sampling made possible by studying gravitationally-lensed sources offers the key to resolving several uncertainties of interpretation of the kinematic and morphological properties of high redshift star-forming galaxies. These include discriminating between single rotating systems and close merging pairs as well as the validity of well-ordered disk models claimed to fit a high proportion of objects studied in the literature. For this study, we also seek to examine possible correlations between the ratio of ordered and random motions (v/σ - typically used as a measure of the extent to which a system is or is not dispersion-dominated) with the presence of a metallicity gradient (see §5), as this might indicate important diagnostics of how disks assemble and mature over a key period in cosmic history. For these and other applications, clearly determining the reliability of the derived values of the rotational velocity and the dispersion is crucially important.

Early LGSAO-based data on lensed sources already highlighted the importance of securing well-sampled data (T. Jones, R. Ellis, et al., 2010; T. A. Jones, Swinbank, et al., 2010). Prior to these studies, it was commonly believed that systemic rotation was confined to the galaxies with larger stellar masses (Genzel, Tacconi, et al., 2006; Law, Steidel, Erb, et al., 2009). However, due to the small number of lensed sources available, limited statistics has remained a problem. T. A. Jones, Swinbank, et al. (2010) presented spatially-resolved dynamics for 6 lensed systems claiming 4/6 showed well-ordered velocity fields - a fraction consistent with that determined

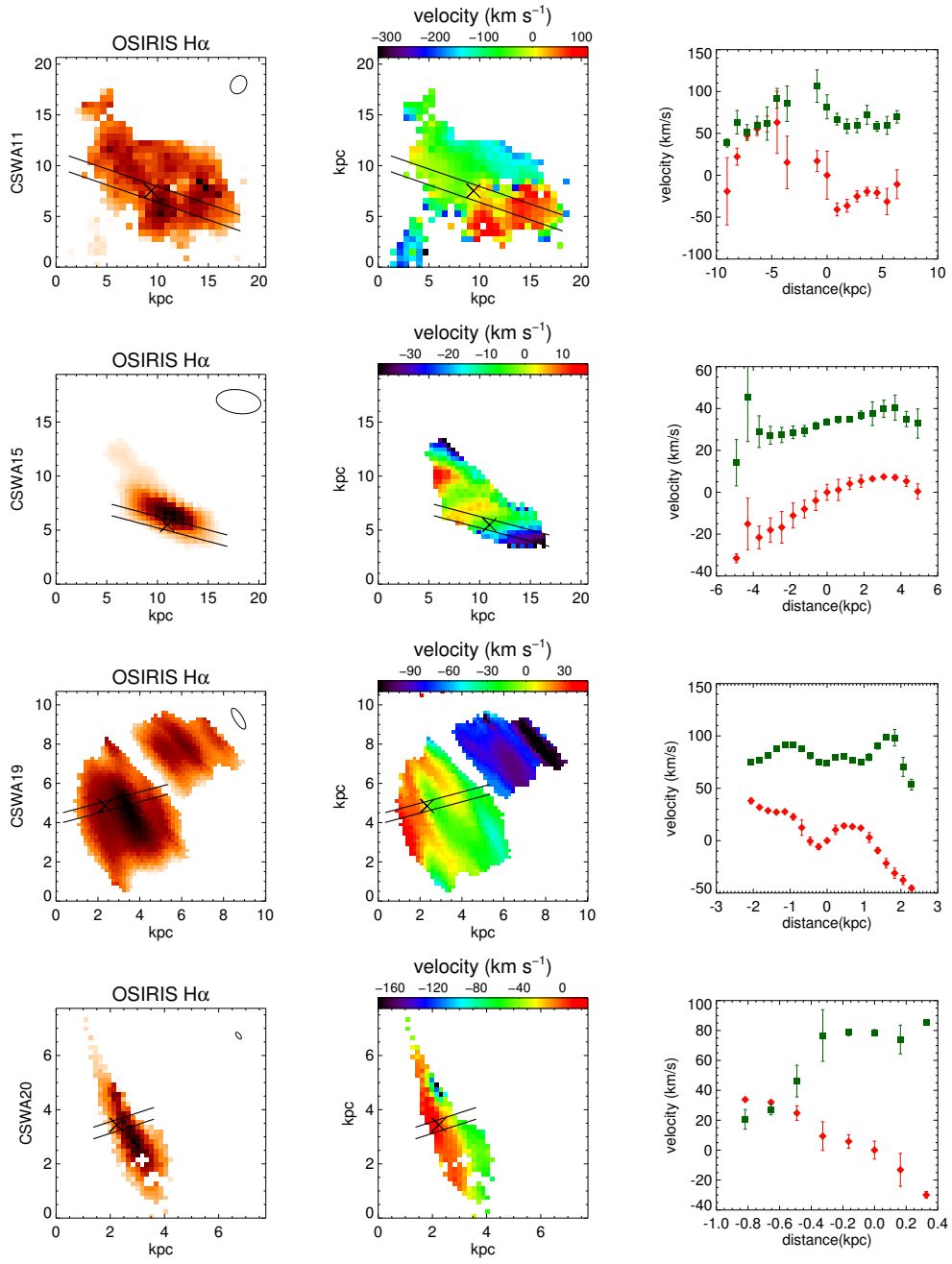

Figure 3.1: From left to right: source plane Hα emission map, two-dimensional velocity field, and one-dimensional velocity (red diamonds) and dispersion (green squares) for each of our targets. Ellipses represent the source-plane PSFs. Solid lines show the slits used to extract the one-dimensional velocity and dispersion with position angles determined from our simple disk model fits (see text in Section 3.4 for details). Black crosses mark the adopted centers of each galaxy.

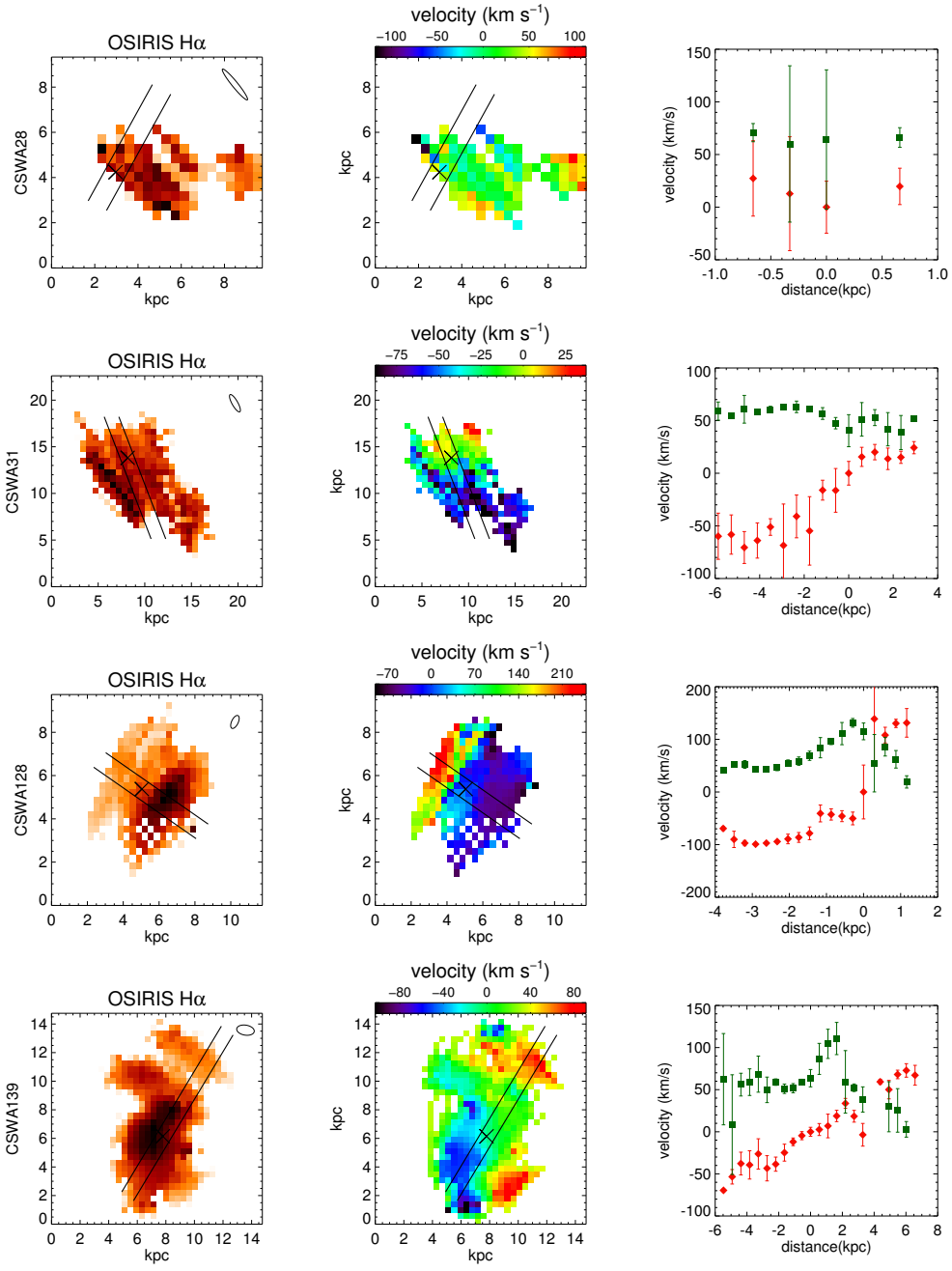

Figure 3.1: continued

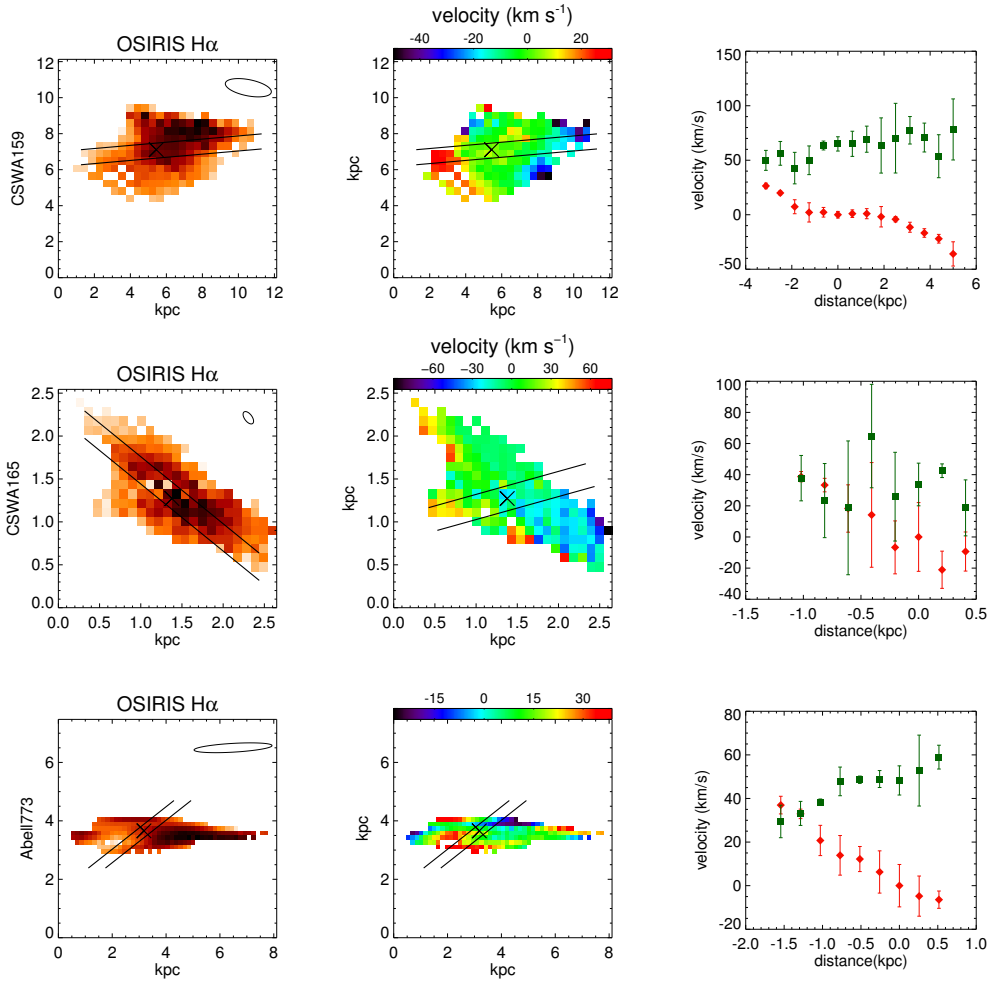

Figure 3.1: continued

for larger, more luminous sources. Recently, Livermore, T. A. Jones, et al. (2015) collated the available data on 17 lensed sources including 10 studied with natural guide star-assisted AO using SINFONI on the VLT and 6 with LGSAO and Keck's OSIRIS, mostly from T. A. Jones, Swinbank, et al. (2010) and T. A. Jones, R. S. Ellis, et al. (2013), similarly concluding ≃60% of these systems with stellar masses in the range $4 \times 10^8 < M/M_\odot < 6 \times 10^{10}$ are consistent with rotating disks. This fraction is lower than that deduced from the KMOS^{3D} survey (Wisnioski, Förster Schreiber, et al., 2015) based on coarser sampled data of a larger (≃30) sample of $1.9 < z < 2.7$ galaxies where, over their entire sample, it is claimed 83% of the sample are rotation-dominated with $v/\sigma > 1$ and at least 70% can be considered 'disk-like' systems.

Our newly-enlarged set of well-sampled resolved spectroscopic data provides a valuable test of the above claims and we now examine both the validity of a simple

disk model as a fit to our source plane velocity fields as well as the extent to which our systems are rotation or dispersion-dominated. Our Hα emission maps and their associated radial velocities fields are presented in Figure 3.1. While our Hα-based radial velocity fields are in many cases indicative of velocity gradients, the associated Hα surface brightness distributions have irregular, asymmetric, and non-disk like shapes. Such irregular emission line morphologies are common at these redshifts and star-formation rates and contrast with more regular distributions in broad-band filters sampling the stellar distributions. To verify this, for each galaxy, we attempted to extract the total continuum spectrum adjacent to the Hα line. Although the resulting signal to noise of the continuum is limited beyond a scale of $\simeq$1 kpc, it provides a valuable indicator of the center of each galaxy.

First, we assess the validity of a rotating disk as a representation of the observed Hα radial velocity field following the method adopted in T. A. Jones, Swinbank, et al. (2010). Those authors used an arctangent function as the simplest model, viz:

$$V(R) = V_0 + \frac{2}{\pi} V_c \arctan \frac{R}{R_t} \tag{3.1}$$

where R is the radius from the disc center,

$$R^2 = (x \cos\theta - y \sin\theta)^2 + (\frac{x \sin\theta + y \cos\theta}{\sin i})^2$$

The simple model has seven unknown parameters: the inclination i, position angle θ, the position of the disk center (α_c, δ_C), the scale radius R_i, an asymptotic velocity V_c, and an overall systemic velocity V_0. The variables x and y represent the distance of each pixel from the disc center. The resulting asymptotic velocity is corrected for the effect of inclination.

We use a Markov-chain Monte Carlo (MCMC) method to find the best fitting disk model parameters. We fit the observed Hα radial velocity map with the model convolved with an elliptical point spread function based on the reconstructed tip-tilt star image. In the MCMC fitting, we set flat priors within the parameter ranges as follows. The disk center is set to lie within the peak of the stellar continuum light distribution (which does not necessarily coincide with the Hα emission peak). The scale radius is confined to lie within the range of the observed Hα map. The asymptotic velocity is $\lesssim 3 \times \Delta V$ where ΔV is the maximum velocity difference along the slit that passes through the disk center for all position angles. We allow the full range for inclination and position angles i.e.,$0° \leq i \leq 90°$ and $0° \leq \theta \leq 360°$. The uncertainty in each derived parameter is calculated from its posterior probability

distribution. The results from these simple disk model fits are summarized in Table 3.3.

Results

Our data reveals many cases where there are significant kinematic deviations from this simple rotating disk model. Because our spatial resolution is <500pc, we find the velocity field for the Hα emission line samples the local bulk motion of gas and is not smeared over a large area as would be the case in coarser-sampled data. Typically our resolution elements have velocity uncertainties of only a few kms sec^{-1} which often results in large reduced χ^2 despite similar r.m.s velocity residuals of $\left\langle v_{\rm model} - v_{\rm data} \right\rangle \sim 30$ km s^{-1} compared to studies sampling $\sim$ kpc scales (e.g. Swinbank, Sobral, et al., 2012).

These significant velocity deviations suggest that several of our systems are undergoing various stages of a merging process. One-dimensional velocity profiles in Figure 3.1 show many galaxies in our sample can be considered as a perturbed disk where the large-scale velocity map exhibits a clear gradient but the velocity dispersion peak is offset from our adopted center. In several cases (e.g. CSWA19, and CSWA139), two distinct Hα sources of different velocities can be discerned, consistent with early stages of a merging pair. There is good evidence from either the velocity profiles or the velocity deviations from our disk model, that CSWA19, 31, 128, 139, and 165 represent merging systems.

At this point, it is interesting to conjecture whether such complex kinematic patterns would have been discernible with a lower spatial resolution. As an experiment, we re-analyzed two systems, CSWA19 and CSWA31 (both clearly poor candidates for pure rotation-dominated systems) as if they had been observed at lower angular resolution. The source plane velocity fields were first smoothed with a typical PSF of 0.6″, rebinned to a coarser resolution of 0.2″, comparable to that achievable with the KMOS^{3D} survey(Wisnioski, Förster Schreiber, et al., 2015), and re-analyzed with our disk model. The results of this experiment are shown in Figure 3.2. In both cases, the χ^2 fit is significantly improved and there is little evidence of any departure from a pure disk model. This accentuates the importance of optimally-sampled data for a robust interpretation of the velocity field.

The second question posed above relates to the precision of the derived rotational velocity and, particularly, the ratio of systemic to random motions which is such an important diagnostic of the maturity of early disk systems. To examine this, we

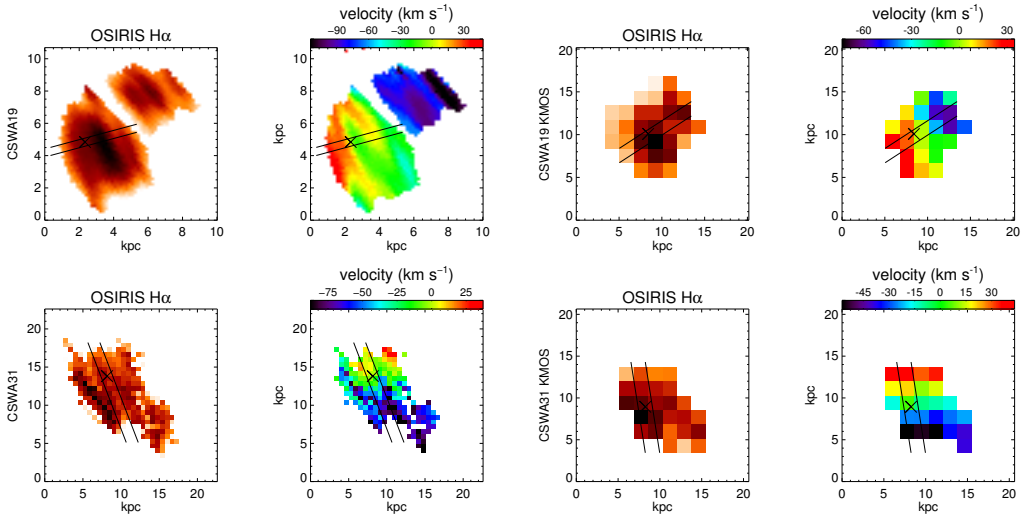

Figure 3.2: The potential dangers of interpreting kinematic data with poor sampling. *Left:* Source plane Hα intensity map and velocity field in this study for CSWA19 (top row) and CSWA31 (bottom row). *Right:* Respective simulated Hα intensity map and velocity field indicating the deterioration in resolution equivalent to that for non-lensed sources studied with KMOS. In both cases, the complex morphology and velocity structure is lost in the poorer sampled data leading to the erroneous conclusion of kinematically well-ordered systems.

first adopt the inclination and position angle values obtained from the disk model fitting. Despite the velocity deviations from the disk model discussed above, the posterior probability distributions of inclination and position angle are generally Gaussian in form suggesting reasonably good estimates of these parameters. Other methods for determining the inclination e.g. analyzing the elliptical shape of the Hα or near-infrared light distributions(e.g. S. F. Newman, Genzel, Förster Schreiber, et al., 2013; Wisnioski, Förster Schreiber, et al., 2015) are less applicable to our less luminous systems with irregular forms. As for the position angle, our disk model fits are generally consistent (within 2σ) with values obtained from a 'peak-to-peak' velocity method where a pseudo-slit placed at the center is rotated to find to yield the largest velocity shear. With the exception in CSWA165 where the discrepancy between the two methods is large and the disk model fit is particularly poor, we adopt the position angle from the peak-to-peak method (θ' in Table 3.3).

Our sample can then be split in two subsets. For five galaxies (CSWA11, 15, 28, A773, 159) where the disk model represents an acceptable fit ($\chi^2_{red} < 20$), the ratio v_c/σ is probably a valid indicator of whether the system is dynamically supported. Adopting the standard often-used threshold $v_c/\sigma \geq 1$ and calculating the intrinsic

velocity dispersion σ from the weighted velocity dispersions of each spaxel along the major axis, all but CSWA15 could be considered to be rotationally-supported systems (see Table 3.3). However, the v_c/σ ratio is clearly inappropriate for the remaining galaxies where the disk model is a much poorer fit. An alternative statistic sometimes used is the observed velocity ratio $\Delta V/2\sigma$ ratio where ΔV is the maximum velocity shear obtained from the peak-to-peak method. Applying the threshold criterion of $\Delta V/2\sigma \geq 0.4$ for rotationally-supported systems adopted by Förster Schreiber, Genzel, Bouché, et al. (2009) to the entire sample, 9 out of 11 of our sources would then be classified as rotation-dominated systems. Figure 3.3 shows the correspondence between these two measures (v_c/σ, $\Delta V/2\sigma$) and the goodness of the disk model (χ^2_{red}) for our sample. Broadly speaking, both above criteria are reasonably consistent in selecting rotationally-supported systems. However, without the highly-sampled data that allows an adequate test of the goodness of fit, a much larger number would be incorrectly placed in this category.

We perform the same kinematic fitting method to the 4 galaxies published in T. A. Jones, R. S. Ellis, et al. (2013) and find consistent categorization using the χ^2_{red} criterion above. T. Jones, R. Ellis, et al. (2010) already performed kinetic model fitting for J0744 and J1206 and we find very similar results. J1148, which T. A. Jones, R. S. Ellis, et al. (2013) considered as a rotating system using a 1-D velocity profile, has a $\chi^2_{red} \approx 6.5$ with a very low velocity residual root-mean-square of 7 km/s suggesting a well-ordered disk. Finally, J1038 which is visibly a merger, has a $\chi^2_{red} \approx 380$ and is categorized as a perturbed disk.

It is illustrative, therefore, to compare our conclusions to the first results of the KMOS^{3D} Survey which reports kinematics result of an IFU survey of ~ 200 galaxies at $0.7 < z < 2.7$ (Wisnioski, Förster Schreiber, et al., 2015). Our galaxy sample has an average $\Delta V/\sin i$ of 148 km s^{-1} and average velocity dispersion of 60 km s^{-1}, both of which are broadly consistent with the range of $\Delta V/\sin i \sim 100 - 200$ and $\sigma \sim 30 - 80$ km s^{-}1 in the KMOS^{3D} data. However, according to Figure 3.3 only 4/11 ($\sim 36\%$) of our sample would be classified as rotationally-supported galaxies via the combination of both criteria in Figure 3.3, whereas in the KMOS^{3D}74% of $z \sim 2$ galaxies were classified as being rotationally-supported. The higher percentages in the KMOS^{3D} analysis may arise at least in part from a combination of the lower spatial resolution where late merging systems are mistaken as a regular rotating system (Figure 3.2) as well as the difficulty of fitting a disk model in such circumstances.

Table 3.3: Kinematic Properties of the Samples

		Simple-disk model results					Peak-to-peak method results			
ID	$\chi^2_{\rm red}$	$\langle V_{\rm model} - V_{\rm data} \rangle_{\rm rms}$	$i(^\circ)$	$\theta(^\circ)$	V_c(km s^{-1})	V_c/σ	ΔV(km s^{-1})	σ(km s^{-1})	$\theta'(^\circ)$	$\Delta V/2\sigma$
cswa11	17	40.2	64^{69}_{56}	16 ± 12	173^{216}_{125}	$1.8^{2.5}_{1.1}$	167 ± 53	94 ± 26	48	0.89 ± 0.19
cswa15	18	12.9	25^{29}_{19}	194 ± 10	29^{40}_{17}	$0.8^{1.2}_{0.5}$	48 ± 3	34 ± 5	205	0.71 ± 0.06
cswa19[a]	9.6	13.6	49^{56}_{43}	155 ± 4	126^{159}_{83}	$1.6^{2.1}_{1.0}$	89 ± 2	78 ± 14	153	0.57 ± 0.05
cswa19[b]	246	42.3	75^{77}_{65}	156 ± 4	377^{442}_{248}	$3.6^{5.8}_{2.8}$	120 ± 7	78 ± 14	150	1.54 ± 0.30
cswa20	35	21.8	40^{47}_{32}	162 ± 2	133^{146}_{99}	$2.2^{3.0}_{1.1}$	46 ± 14	62 ± 25	168	0.37 ± 0.09
cswa28	6.6	25.3	23^{47}_{18}	129 ± 10	91^{116}_{49}	$1.6^{2.2}_{0.8}$	29 ± 25	56 ± 12	127	0.26 ± 0.12
cswa31	93	20.8	22^{27}_{19}	249 ± 3	122^{135}_{110}	$2.2^{2.6}_{1.8}$	102 ± 12	56 ± 8	244	0.91 ± 0.08
cswa128	990	69.9	57 ± 3	210 ± 2	570^{578}_{554}	$8.6^{12.4}_{4.8}$	262 ± 8	66 ± 29	208	1.98 ± 0.44
cswa139	38	25.6	24 ± 4	297 ± 15	265^{290}_{211}	$4.3^{6.5}_{1.9}$	166 ± 27	62 ± 32	294	1.34 ± 0.36
cswa159	7.6	12.7	13^{19}_{10}	175 ± 9	141^{191}_{85}	$2.3^{3.5}_{1.1}$	79 ± 29	61 ± 22	164	0.65 ± 0.17
cswa165	141	17.3	57 ± 5	164 ± 5	359^{384}_{307}	$5.9^{8.5}_{3.1}$	83 ± 28	44 ± 19	218	0.94 ± 0.26
a773	17	53.7	22^{37}_{17}	144 ± 12	176^{225}_{108}	$3.9^{5.3}_{2.2}$	54 ± 22	45 ± 10	158	0.60 ± 0.28

[a] Results from fitting to the main galaxy only.
[b] Results from fitting to both mergers together as one system.

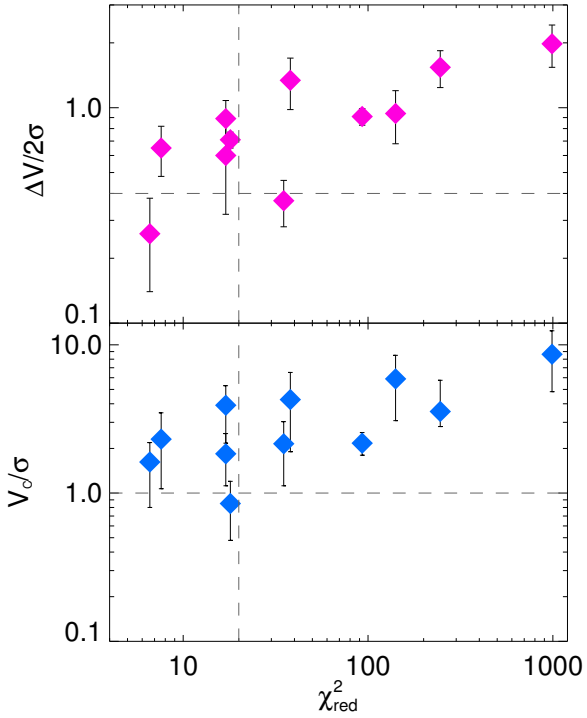

Figure 3.3: A comparison of two methods for determining the degree of rotational support in our sample plotted against the goodness-of-fit in the disk model. v_c/σ from our disk fits and $\Delta V/2\sigma$ from the peak-to-peak method is considered in the context of the reduced χ^2. Horizontal dashed lines represent the usually adopted thresholds, $\Delta V/2\sigma \geq 0.4$ and $v_c/\sigma > 1$ for rotationally-supported systems and the vertical dashed line represents $\chi^2 = 20$, below which we consider the disk model to be appropriate. Without adequate sampling, a larger fraction of our sample would be incorrectly considered to be rotationally-supported.

3.5 Metal Gradients

Given the capability of the current integral field instruments, the only practical method to measure metallicity in high-redshift galaxies is to use strong-line metallicity calibrators. The direct measurement of metallicity (T_e method) is not feasible for individual galaxies at this redshift since it relies on the measurement of emission lines that are too faint to detect. In this work, we use the [N II] λ6584/Hα ratio (N2) and the ([O III] λ5008/Hβ)/([N II] λ6584/Hα) ratio (O3N2) with the calibrations from Pettini et al. (2004) - PP04, Steidel, Rudie, et al. (2014) - S14, and Maiolino, Nagao, et al. (2008) - M08:

$$12 + \log(\mathrm{O/H}) = 8.90 + 0.57 \times \mathrm{N2} \qquad \text{(N2, PP04)}$$

$$12 + \log(\mathrm{O/H}) = 8.62 + 0.36 \times \mathrm{N2} \qquad \text{(N2, S14)}$$

Table 3.4: Metallicity Gradients

ID	[N II]/Hα	Central 12+log(O/H) (from N2 PP04)	N2 PP04 dex kpc^{-1}	N2 S+14 dex kpc^{-1}	N2 M+08 dex kpc^{-1}	O3N2PP04 dex kpc^{-1}	O3N2 S+14 dex kpc^{-1}	O3N2 M+08 dex kpc^{-1}
Objects classified as rotationally-supported systems ($\chi^2 < 20$)								
cswa11	0.14 ± 0.01	8.54 ± 0.06	−0.07 ± 0.02	−0.05 ± 0.01	−0.10 ± 0.03	–	–	–
		8.60 ± 0.06	−0.11 ± 0.02	−0.07 ± 0.01	−0.16 ± 0.02	–	–	–
cswa15	0.03 ± 0.01	8.16 ± 0.01	−0.04 ± 0.01	−0.02 ± 0.01	−0.05 ± 0.02	−0.02 ± 0.01	−0.02 ± 0.01	−0.01 ± 0.01
		8.17 ± 0.01	−0.03 ± 0.01	−0.02 ± 0.01	−0.05 ± 0.01	−0.02 ± 0.01	−0.01 ± 0.01	−0.00 ± 0.01
cswa28	0.09 ± 0.01	8.10 ± 0.01	0.11 ± 0.03	0.07 ± 0.24	0.17 ± 0.59	–	–	–
		8.33 ± 0.12	0.04 ± 0.03	0.02 ± 0.01	0.07 ± 0.04	–	–	–
cswa159	0.08 ± 0.01	8.50 ± 0.01	−0.01 ± 0.01	−0.01 ± 0.01	−0.03 ± 0.02	0.00 ± 0.01	0.00 ± 0.01	−0.01 ± 0.01
		8.49 ± 0.03	−0.01 ± 0.01	0.00 ± 0.01	−0.02 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	−0.01 ± 0.01
a773	0.17 ± 0.07	8.42 ± 0.10	−0.13 ± 0.18	−0.08 ± 0.06	−0.19 ± 0.28	−0.25 ± 0.41	−0.22 ± 0.35	0.01 ± 0.30
		8.43 ± 0.05	−0.05 ± 0.02	−0.03 ± 0.01	−0.07 ± 0.03	−0.03 ± 0.03	−0.02 ± 0.02	−0.08 ± 0.02
Objects with $\chi^2 > 20$								
cswa19	0.09 ± 0.01	8.30 ± 0.01	−0.01 ± 0.01	0.01 ± 0.01	−0.01 ± 0.07	0.12 ± 0.04	0.10 ± 0.01	0.07 ± 0.05
		8.26 ± 0.01	0.02 ± 0.01	0.01 ± 0.01	0.02 ± 0.01	0.01 ± 0.01	0.01± 0.01	0.03 ± 0.01
cswa20	0.05 ± 0.01	7.89 ± 0.02	−0.15 ± 0.05	−0.10 ± 0.07	−0.29 ± 0.45	−0.41 ± 0.19	−0.36 ± 0.17	−1.00 ± 0.24
		7.97 ± 0.04	0.05 ± 0.02	0.03 ± 0.01	0.09 ± 0.03	0.03 ± 0.02	0.02 ± 0.02	0.25 ± 0.05
cswa31	0.33 ± 0.06	8.50 ± 0.03	0.02 ± 0.01	0.01 ± 0.01	0.03 ± 0.01	–	–	–
		8.54 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	–	–	–
cswa128	0.10 ± 0.02	8.57 ± 0.03	−0.10 ± 0.02	−0.06 ± 0.01	−0.12 ± 0.02	−0.09 ± 0.02	−0.08 ± 0.02	−0.15 ± 0.02
		8.51 ± 0.01	−0.04 ± 0.01	−0.02 ± 0.01	−0.05 ± 0.01	−0.04 ± 0.01	−0.03 ± 0.01	−0.06 ± 0.01
cswa139	0.17 ± 0.08	8.25 ± 0.01	−0.01 ± 0.01	−0.01 ± 0.01	−0.02 ± 0.02	0.03 ± 0.01	0.03 ± 0.01	0.02 ± 0.01
		8.12 ± 0.03	0.04 ± 0.01	0.02 ± 0.01	0.05 ± 0.01	0.02 ± 0.01	0.01 ± 0.01	0.04 ± 0.01
cswa165	0.26 ± 0.06	8.53 ± 0.01	0.03 ± 0.01	0.02 ± 0.02	0.04 ± 0.05	0.13 ± 0.01	0.11 ± 0.01	0.16 ± 0.05
		8.53 ± 0.01	0.01 ± 0.02	0.01 ± 0.01	0.01 ± 0.04	0.01 ± 0.02	0.00 ± 0.01	0.13 ± 0.03

For each object, the top line contains values derived along each galaxy's 'major axis'. The bottom line contains values derived from radial binning.

$$12 + \log(\mathrm{O/H}) = 8.73 - 0.32 \times \mathrm{O3N2} \qquad \text{(O3N2, PP04)}$$

$$12 + \log(\mathrm{O/H}) = 8.66 - 0.28 \times \mathrm{O3N2} \qquad \text{(O3N2, S14)}$$

where N2$\equiv \log([\mathrm{N\ II}]\lambda6584/\mathrm{H}\alpha)$ and O3N2$\equiv \log(([\mathrm{O\ III}]\ \lambda5008/\mathrm{H}\beta)/([\mathrm{N\ II}]\lambda6584/\mathrm{H}\alpha))$.

The PP04 relations were primarily calibrated from the T_e method based on nearby extragalactic HII regions. The M08 relations were based on local galaxies that cover a wider range of metallicity compared to the PP04 relations. The S14 relations were based on a subset of the same nearby extragalactic HII regions sample in PP04 relations that are limited to the range of N2 and O3N2 observed in $z \sim 2.3$ sample which have shown to lower the systematic offset between the two indicators when applied to high redshift sample. To counter the possible bias of the inferred metallicity due to the higher N/O ratios in high redshift galaxies than the ratios in local galaxies, we will compare our measured gradients from the N2 calibrators to the gradients from the O3N2 indicators which have been shown to be less dependent on the N/O ratio than the N2 indicator (Steidel, Rudie, et al., 2014).

For each calibrator, we calculate the metallicity gradient using two methods: from the pixels along the major axis and from all pixels binned radially. The galaxy center, position angle, inclination, and pseudo slit along major axis are the same as those in Section 3.4. An advantage of the metallicity gradient measured from the pixels along the major axis is that it is independent of the inclination and thus less model-dependent. However, it is based on fewer measurements than the metallicity gradient measured from all pixels and sometimes not feasible when most of the HII regions do not lie along the major axis (e.g. CSWA20). The best-fit gradients and central metallicities are shown in Table 3.4 and plots are shown in Figure 3.9.

In detail, we calculate an average N2 ratio (or O3N2 ratio) for each radial bin along the major axis (or along the annuli) for PP04 and S14 calibrators. We assume that metallicity 12 + log (O/H) is a linear function of radius. Hence, the line ratio is an exponential function of radius, metallicity gradient, and central metallicity. We fit the N2 ratios (or O3N2 ratios) with the weighted least-square regression to find the metallicity gradient and central metallicity for each galaxy.

For the M08 calibrator, we calculate the N2 ratio (and [O III]/Hβ ratio, if possible) for each radial bin and find the metallicity for that radial bin with a brute force maximum likelihood estimation. We then obtain the metallicity gradient and the central metallicity with a linear fit.

We check the consistency between the metallicity gradients derived from pixels along major axis and radial annuli obtained in section 3.4. We found that the results from the two methods are consistent with each other within one standard deviation especially when the rotation is relatively well-described by a simple disk rotation. The deviation is larger when the major axis is short and the measurement along the major axis suffers from a poor sampling e.g. CSWA20 and Abell773.

Metallicity gradients derived from the O3N2 calibrators are consistent with the gradients derived from the N2 calibrators when using radial binning method and comparing between the same set of calibrators e.g. PP04 N2 with PP04 O3N2. Although O3N2 calibrators give more accurate metallicity measurement than N2 calibrators, the consistency in the gradients derived from the two calibrators suggests that the metallicity gradients measured with N2 calibrators should not be heavily affected by a possible variation in N/O ratios in high redshift galaxies. Hence, for the three galaxies that we do not have information on [O III]/Hβ, the derived metallicity gradients should be relatively reliable. When compared between the S14 and PP04 measurements, the metallicity gradients derived with S14 calibrators are typically flatter than those derived with PP04 calibrators.

In further analysis, we use the metallicity gradients derived from PP04 in radial binning because we can avoid the bias from poor sampling in short major axis cases. The PP04 calibrator is chosen so that we can easily compare with other observed gradients from previous observations.

3.6 AGN Contamination and Bias from Low Signal to Noise Data

We now examine whether our sample is affected by line emission from AGN. Since our metallicity calibrators are only applicable to bounded HII regions ionized by hot stars, any emission from nuclear activity would give misleading results for central metallicities and gradients. Furthermore, the presence or absence of AGN is interesting in terms of understanding the relationship between growth of galaxies and their central supermassive black holes. Optical spectra of AGN are characterized by high ratios of collisional to Balmer lines, and by broad line widths arising from outflows and/or broad line region kinematics. These signatures have been confirmed in the nuclear regions of many galaxies at $z \simeq 1-2$ with spatially resolved spectroscopy (Wright et al., 2010; S. F. Newman, Buschkamp, et al., 2014; Förster Schreiber, Genzel, S. F. Newman, et al., 2014; Genzel, Förster Schreiber, et al., 2014). Notably, Genzel, Förster Schreiber, et al. (2014) show that the fraction

of galaxies showing signatures of AGN is strongly correlated with stellar mass.

For each of the eight new galaxies in this paper with [O III]/Hβ information, we construct a "BPT":[N II]/Hα and [O III]/Hβ diagram from the spatially resolved spectra. Most show a distribution of values entirely consistent with star-forming regions as seen in other $z \sim 2$ galaxies (Steidel, Rudie, et al., 2014; Shapley, Reddy, et al., 2015). We thus only show two examples where the interpretation is less clear (Figure 3.4). CSWA15 shows a central region in whose strong line ratios significantly lie at an average of 0.16 ± 0.06 dex above the maximum starburst classification line (Kewley, Dopita, et al., 2001) as shown in the upper panel of Figure 3.4. However, the central [N II]/H$\alpha = 0.03 \pm 0.01$ and [S II]$\lambda\lambda 6716, 6731$/H$\alpha = 0.13 \pm 0.03$ are both low and atypical of AGN. These ratios are at the extreme low end probed by current $z \simeq 2$ samples (e.g. Steidel, Rudie, et al., 2014) and broadly consistent with an extension of the star-forming locus seen in these surveys. Alternatively, the line ratios are consistent with theoretical expectations for an AGN with gas metallicity of $\simeq 0.5$ solar (Groves et al., 2006), but the surrounding (and clearly star-forming) regions of CSWA15 indicate significantly lower metallicity. If the central line ratios are affected by AGN then the gas-phase metallicity gradient would be significantly steeper than we infer assuming stellar ionization.

As an additional check for nuclear activity, we fit Hα from the central regions of CSWA15 with a double Gaussian profile. The best-fit broad component has a line width $\sigma = 153$ km s^{-1} and comprises 29% of the total flux (c.f. $\sigma = 59$ km s^{-1} for the dominant narrow component). This is similar to stacked spectra of outer disk regions in $z \simeq 2$ galaxies ($\sigma \sim 200$ km s^{-1} and ~40% of flux in the broad component; S. F. Newman, Genzel, Förster Schreiber, et al. 2013, Genzel, Förster Schreiber, et al. 2014) and is not indicative of AGN-driven outflows ($\sigma \gtrsim 500$ km s^{-1}) or broad-line regions ($\sigma \gtrsim 1000$ km s^{-1}). The spectrum of CSWA15 is therefore consistent with expectations for emission from star formation and associated outflows, and does not show strong evidence of an AGN.

The incidence of AGN in our sample is consistent with expectations given their stellar masses. We find no clear signatures of AGN and only one possible case discussed above out of the 12 galaxies with available BPT diagnostics, including those from T. A. Jones, R. S. Ellis, et al. (2013). An example of typical BPT diagrams that are compliant with star-forming regions in our sample, CSWA165, is shown in the lower panel of Figure 3.4. The remaining galaxies in our sample have low central [N II]/Hα and [S II]/Hα ratios indicating no evidence for AGN. Genzel,

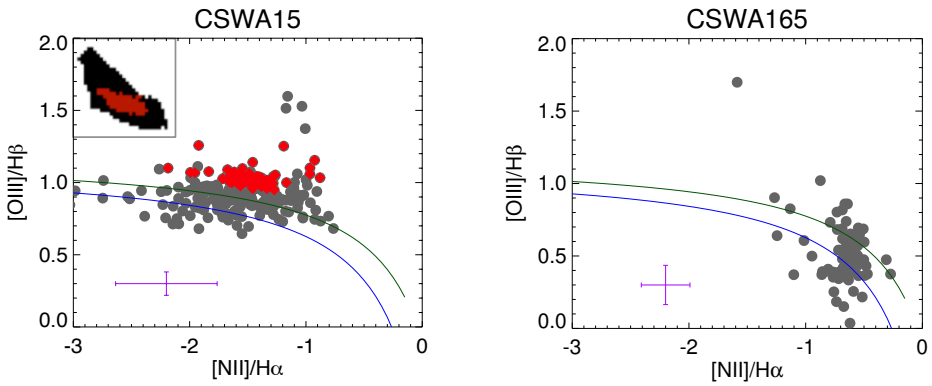

Figure 3.4: BPT diagrams for CSWA15 (left) and CSWA165 (right). Gray points are from individual pixels in each galaxy. Red points are from pixels whose values lie significantly above the maximum starburst line (green curve, Kewley, Dopita, et al. (2001)). The corresponding region to the red points is shown in red in the map at the corner. Blue curve is the best fitted line for the $z \sim 2.3$ samples in Steidel, Rudie, et al. (2014). Median values in uncertainties are shown in the purple error bars. A few pixels have large error bars e.g. the data point at the top of the lower plot.

Förster Schreiber, et al. (2014) find similar results using the same methods: zero secure and only two potential AGN in a sample of 17 $z \simeq 1$–2 galaxies with stellar mass $\log M_*/M_\odot = 9.4$–10.3, comparable to that of our sample. They likewise find an AGN fraction $< 10\%$ at $\log M_*/M_\odot < 10.3$ based on independent X-ray, infrared, and radio indicators. Therefore, the absence of AGN in our sample is not unexpected given the stellar mass range probed.

We then consider the possible bias in measured metallicity gradients due to possibly low signal-to-noise ratio of [N II] emission lines. We select CSWA128 to be a fiducial sample in this analysis due to its moderately steep gradient so that we can examine possible degradation of this gradient as we increase the noise. We adopt the 2σ detection limit of the [N II] emission line for CSWA128;the limits for other galaxies are similar. We then synthetically add Gaussian noise to the data cube up to twice the detection limit and re-measure the metallicity gradient with the same procedure. We compare the measured gradients when the noise is added to the data cube to the gradients measured from the original data cube in Figure 3.5. We find that the gradients measured from the noise-added data cubes are consistent with the original measurement up to doubling the noise or a quarter of the exposure time for CSWA128. Accordingly, we can be confident that the measured gradients in other galaxies are not likely to be affected given our [N II] detection limits.

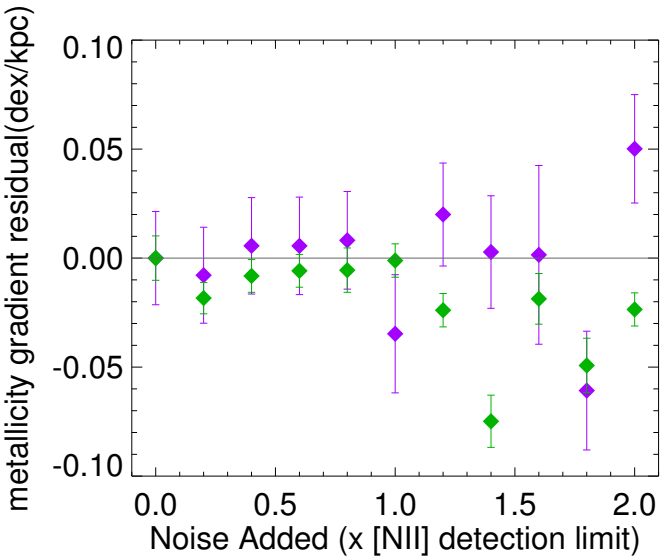

Figure 3.5: Verifying the validity of our derived metal gradients for low signal/noise [N II] data. Metallicity gradient residuals are shown for CSWA128 as a function of the amount of synthetic noise added (see text in Section 3.6 for technical details). Gradients are derived with PP04 method both along the galaxy 'major axis' (purple) and using radial binning(green).

3.7 Discussion

Previous observations (Yuan, Kewley, Swinbank, et al., 2011; T. A. Jones, R. S. Ellis, et al., 2013; T. Jones, Wang, et al., 2015) led to a simple picture whereby isolated galaxies at high redshifts tend to have steep metallicity gradients which gradually flatten with cosmic time, while interacting galaxies have no discernible gradients at all. Conceptually this could be understood if metal gradients only became properly established when galaxies are kinematically well-ordered, with gradients flattening as galaxies grow in size. However, the enlarged data set presented here shows a large diversity of gradients which is incompatible with this picture. The redshift-dependent behavior of metallicity gradients in our sample is summarized in Figure 3.6. As in previous work we find (negatively) steeper gradients in rotationally supported systems compared to the merging and dynamically immature systems. However, the gradients in rotationally supported galaxies are much flatter (> -0.1 dex/kpc) than those previously observed in T. A. Jones, R. S. Ellis, et al. (2013) (~ -0.3 dex/kpc), indicating less evolution in the gradient slope at $z < 2$. We now seek to explain the diversity in observed gradient slopes especially among isolated rotating galaxies.

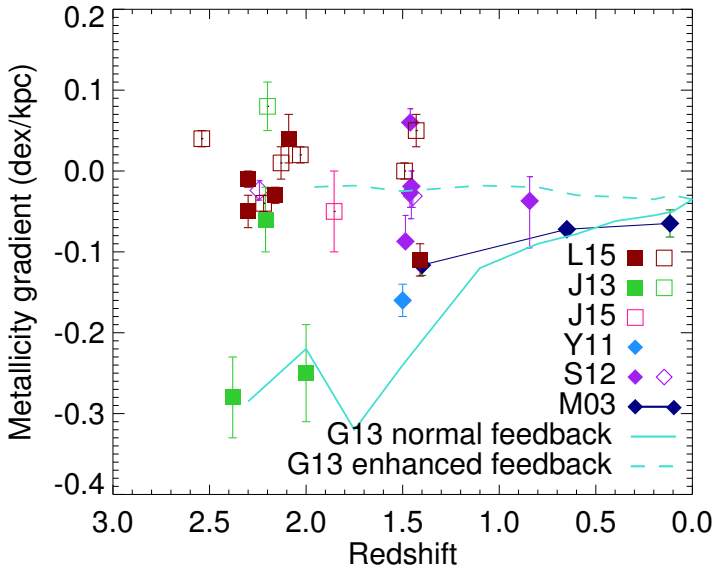

Figure 3.6: The evolution of metallicity gradient with redshift. Red squares represent lensed galaxies from the present study (L15). Green squares are the four lensed galaxies in T. A. Jones, R. S. Ellis, et al. (2013, J13). Other data points are measured from a lensed starburst dwarf (T. Jones, Wang, et al., 2015, J15), a lensed galaxy (Yuan, Kewley, Swinbank, et al., 2011, Y11), non-lensed galaxies observed with adaptive optics (Swinbank, Sobral, et al., 2012, S12), and Milky Way's planetary nebulae (Maciel et al., 2003, p. M03). Open and filled symbols represent kinematically well-ordered and disturbed systems respectively. Turquoise lines show predictions from the hydrodynamical simulations of Gibson et al. (2013, G13) emphasizing the strong sensitivity to the incorporated feedback.

One possibility for the difference seen between this work and that of T. A. Jones, R. S. Ellis, et al. (2013) could be procedural. In the present sample, we avoided a possible selection bias towards unusually active or metal-rich sources. In T. A. Jones, R. S. Ellis, et al. (2013), integrated [N II] and Hα fluxes were first measured for a larger sample with a slit spectrograph and only those four sources with relatively bright emission lines (particularly [N II]) were selected for subsequent study with OSIRIS. Following upgrades to the OSIRIS grating and Keck I adaptive optics system, this pre-screening step was avoided in the current sample (with the exception of CSWA31). To test for this bias, the left panel of Figure 3.7 correlates the metallicity gradient with the integrated [N II]/Hα line ratio for both isolated or rotationally-supported and kinematically disturbed systems. All galaxies with [N II]/Hα < 0.1 have flat gradients (< 0.1 dex kpc^{-1}). Moreover, for galaxies with [N II]/Hα > 0.1, there is a divergence between isolated or rotationally-supported and merging or

dynamically-immature systems with the latter showing zero gradients irrespective of the metallicity. The right panel of Figure 3.7 examines the dependence on the integrated star formation rate to check whether the T. A. Jones, R. S. Ellis, et al. (2013) sample is biased to more intense star-forming systems; however, no strong trend is revealed.

A more likely possibility for the diversity of gradients at $z \simeq 2$–2.5 is variation in feedback strength. Numerical simulations suggest that metallicity gradients are highly sensitive to feedback in the form of outflows and "galactic fountains" (Pilkington et al., 2012; Gibson et al., 2013; Anglés-Alcázar et al., 2014), in the sense that stronger feedback (i.e., higher mass loading factors) leads to more gas mixing and therefore flatter gradients. This is illustrated in Figure 3.6 which shows evolution of the metallicity gradient in a simulated galaxy using two different feedback prescriptions. Simple analytical chemical evolution models show that stronger feedback also results in lower gas-phase metallicity (e.g. T. A. Jones, R. S. Ellis, et al., 2013). The correlation of gradient slope with integrated metallicity, as shown in the left panel of Figure 3.7, could therefore also be due to a variation in feedback strength. In fact, the enhanced feedback scheme in Figure 3.6 was preferred over the normal feedback scheme considering the stellar mass-halo mass relationship matching (Stinson et al., 2013).

While Figure 3.7 provides evidence that gradient slopes are affected to different degrees by feedback, the trend has a large scatter suggesting that other factors are likely important. One obvious possibility is the degree of rotational support compared to random motions which would mix the gas and flatten any gradients. This effect is clearly seen in the sense that merging and interacting systems have flatter gradients compared to isolated galaxies. Among isolated galaxies, we would naively expect those with higher ratios of rotational to random motion (i.e., higher v/σ) to have stronger gradients in metallicity and perhaps other properties. However, Figure 3.8 shows very little dependence on v/σ. We compare our observational results with $z = 2$ galaxies drawn from the Illustris simulation (Vogelsberger et al., 2014; D. Nelson et al., 2015)[2], a large-volume cosmological hydrodynamical simulation with baryonic feedback. Metallicity gradients in the simulation data were constructed based on a linear fit to all of the star forming gas within a radius of 10 kpc from each galaxyÕs center. The simulated galaxies have gas-phase kinematics and negative metallicity gradients which are similar to our observations

[2]Illustris data is available through `http://www.illustris-project.org`

and likewise show no clear trend between gradient slope and rotational support. In essence, galaxies with gas disks supported by rotation have radial metallicity gradients which are similar to systems dominated by random motions, in contrast to expectations that bulk velocity dispersion should flatten any gradients. This suggests that gradients are established on timescales of the order of a dynamical time, sufficiently short that prominent random motions are not effective at mixing the gas.

To summarize, we find a large diversity of metallicity gradients among rotating galaxies at $z \simeq 2$. Most rotating galaxies in our sample have negative gradients, while interacting systems have flatter gradients, as is found in local galaxy samples. Surprisingly, we find that the degree of ordered rotation versus random bulk motion has no discernible effect on the gradient slope even among isolated galaxies, indicating that gradients are formed and destroyed on timescales comparable to the galaxy dynamical time ($\sim 10^8 - 10^9$ yr). The steepest observed gradients with slopes ~ -0.3 dex/kpc are less common than found in our previous work suggesting that this is a relatively rare or short-lived phase at high redshift. It would therefore be interesting to verify the time-dependence of steep gradients in the numerical simulations. Finally, we find a weak correlation between galaxy-integrated metallicity and gradient slope which we interpret as an effect of feedback: the same feedback which reduces the gas-phase metallicity evidently also flattens the gradients, as found in cosmological zoom-in simulations. The diversity of metallicity gradients at $z \simeq 2$ is therefore likely caused in part by differences in the recent feedback history. As we argued in T. A. Jones, R. S. Ellis, et al., 2013, the evolution of metallicity gradients is evidently a sensitive probe of feedback in the form of gaseous outflows and galactic fountains.

3.8 Summary

We present spatially resolved kinematics and gas-phase metallicity gradient measurements for a total of 15 gravitationally lensed star-forming galaxies at $z \sim 2$ based on the analysis of strong emission lines of Hα, [N II], [O III], and Hβ. 11 new sources were observed with the Keck laser-assisted adaptive optics systems and the upgraded OSIRIS integral field unit spectrograph to which we add earlier data on four sources presented in T. A. Jones, R. S. Ellis, et al. (2013). With the aid of gravitational lensing, the typical source plane spatial resolution for each source is < 500 pc which is considerably better than for other systems studied at similar redshifts.

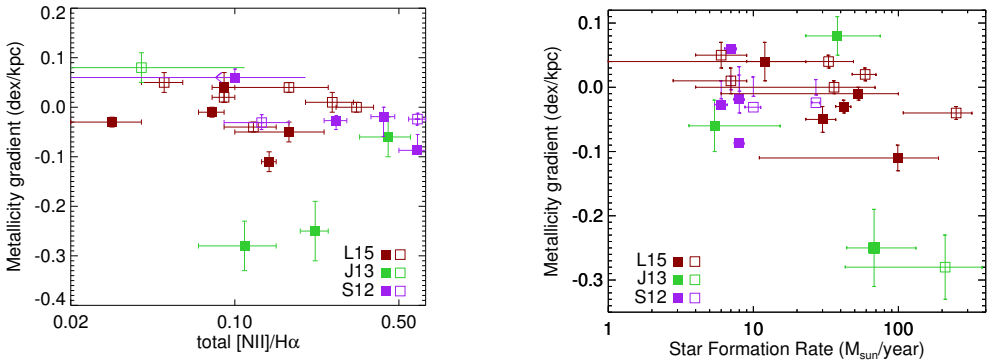

Figure 3.7: Correlation between the observed metallicity gradient for the present sample (L15), T. A. Jones, R. S. Ellis, et al. (2013, J13) and Swinbank, Sobral, et al. (2012, S12) and the integrated [N II]/Hα ratio (left panel) and the star formation rate (right panel). Filled symbols represent isolated galaxies; open symbols represent interacting and/or kinematically-immature galaxies.

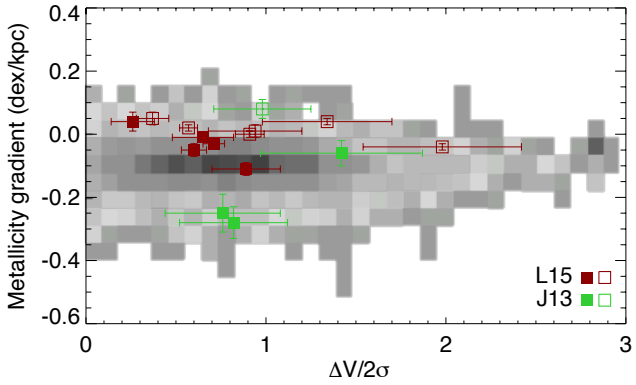

Figure 3.8: The absence of a correlation between the metallicity gradient and the degree of rotational support. Data points represent the present sample and the grey scale represent the output of the Illustris simulation for a sample of $z \simeq 2$ galaxies (Vogelsberger et al., 2014; D. Nelson et al., 2015).

We found the following key results:

- High spatial resolution is crucial in diagnosing the kinematic properties and dynamical maturity of $z \simeq 2$ galaxies. We compare our observed velocity map of interacting pairs of merging galaxies with those simulated at the lower resolution typical in non-lensed surveys and demonstrate how easily such sources can be mistaken to represent rotationally-supported disks in poorer sampled data. Even for galaxies with no clear morphological sign of interaction, we often find significant large deviations in the velocity field compared to rotating disk models which would not be apparent in seeing-limited data. As a result we observe a significantly lower fraction (36%) of rotationally-supported systems in our sample than has been claimed (≃74%) from larger kinematic surveys undertaken with lower spatial resolution.

- We find a much higher fraction of $z \simeq 2$ galaxies have weak or non-existent metallicity gradients than in previous studies of smaller samples observed in the same redshift range. It seems unlikely that such a change arises as the result of a bias in the earlier sample which pre-selected sources with stronger [N II] lines and hence increased metallicity. We observe only a weak correlation between the presence of a gradient and the metallicity and none with the degree of rotational support, the latter being consistent with the recent predictions of the Illustris hydrodynamical simulations. We argue that variations in gas and metal mixing due to feedback most likely play the dominant role in modifying metal gradients and thus can explain the sizable scatter we see in our enlarged sample. The sensitivity of the observed metal gradient to the various modes of feedback indicates it will remain a promising tool for understanding galaxy assembly.

The authors recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. We acknowledge useful discussions with Chuck Steidel, Phil Hopkins, Xiangcheng Ma, Drew Newman and Paul Torrey, and thank Paul Torrey for his Illustris metal gradient predictions tailored to the context of our observations. Support for AZ was provided by NASA through Hubble Fellowship grant #HST-HF2-51334.001-A awarded by STScI. TAJ acknowledges support from NASA through Hubble Fellowship grant HST-HF2-51359.001-A awarded by the

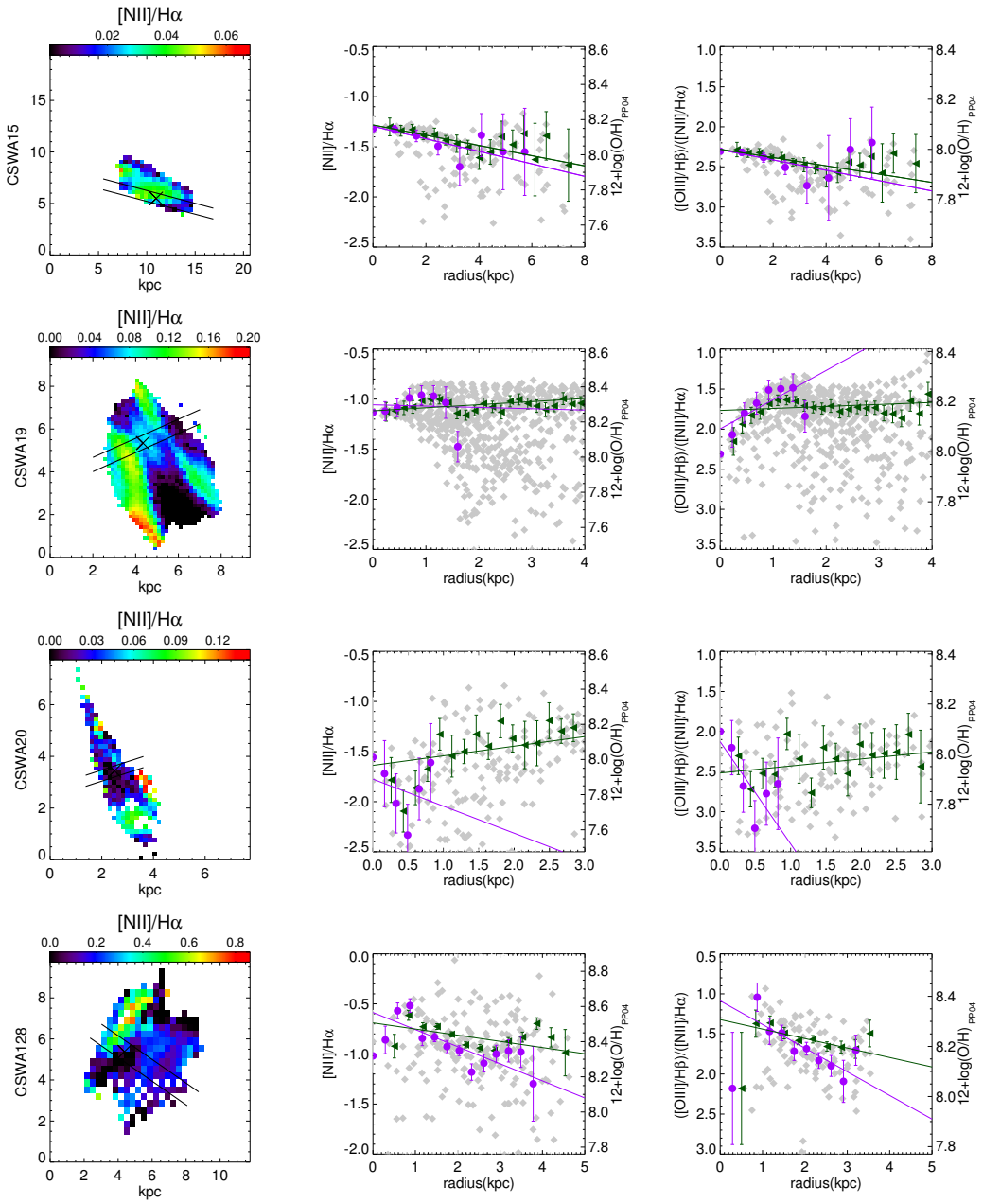

Figure 3.9: From left to right: source plane [N II]/Hα map, radial metallicity gradient from N2 calibrators, and radial metallicity gradient from O3N2 calibrators for subsample of our targets. The gray dots are measurements from each source plane pixels. Green points and lines are measurements from radial binning. Purple points and lines are measurements along 'major axis'.

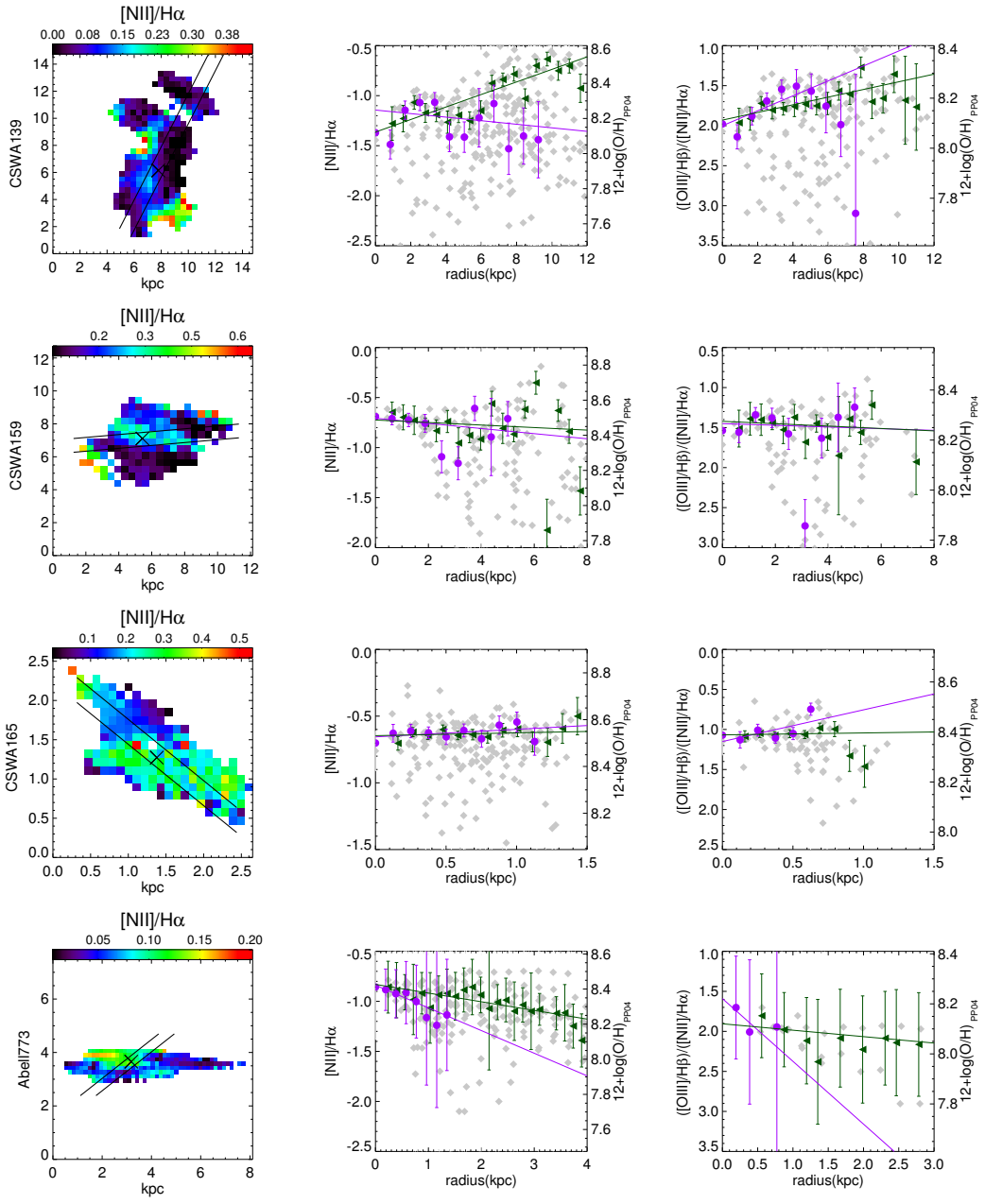

Figure 3.9: continued

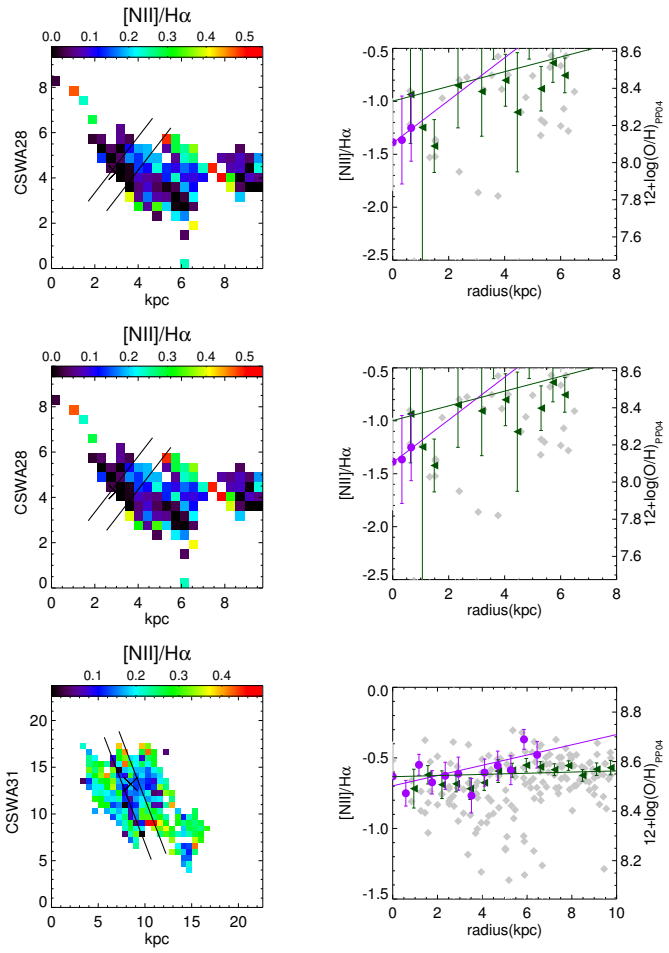

Figure 3.9: continued

Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555

Chapter 4

EVOLUTION OF THE STELLAR MASS–METALLICITY RELATION - I: GALAXIES IN THE Z ∼ 0.4 CLUSTER CL0024

Leethochawalit, N., Kirby, E. N., Moran, S. M., Ellis, R. S., & Treu, T. (2018), ApJ 856, 15, 15

Abstract

We present the stellar mass–stellar metallicity relationship (MZR) in the Cl0024+1654 galaxy cluster at $z \sim 0.4$ using full spectrum stellar population synthesis modeling of individual quiescent galaxies. The lower limit of our stellar mass range is $M_* = 10^{9.7} M_\odot$, the lowest galaxy mass at which individual stellar metallicity has been measured beyond the local universe. We report a detection of an evolution of the stellar MZR with observed redshift at 0.037 ± 0.007 dex per Gyr, consistent with the predictions from hydrodynamical simulations. Additionally, we find that the evolution of the stellar MZR with observed redshift can be explained by an evolution of the stellar MZR with their formation time, i.e., when the single stellar population (SSP)- equivalent ages of galaxies are taken into account. This behavior is consistent with stars forming out of gas that also has an MZR with a normalization that decreases with redshift. Lastly, we find that over the observed mass range, the MZR can be described by a linear function with a shallow slope, ([Fe/H] $\propto (0.16 \pm 0.03) \log M_*$). The slope suggests that galaxy feedback, in terms of mass-loading factor, might be mass-independent over the observed mass and redshift range.

4.1 Introduction

One of the best-known properties in observed galaxies is the tight correlation between galaxy stellar mass and gas-phase metallicity, i.e., the mass–metallicity relation (MZR). Several large galaxy surveys, such as the Sloan Digital Sky Survey (SDSS), have confirmed that galaxies at all redshifts with higher stellar masses retain more metals than galaxies with lower stellar masses (e.g. Tremonti et al., 2004; Sanders et al., 2015; Guo et al., 2016; Onodera, Carollo, S. Lilly, et al., 2016). While the details of the evolution of the gas-phase MZR over redshift are still debated, mainly due to the different metallicity indicators (e.g. Steidel, Rudie, et al.,

2014; Kewley, Zahid, et al., 2015; Strom et al., 2017; Bian et al., 2017), galaxies at higher redshifts generally follow the same trend as the local MZR but are somewhat offset toward lower metallicities (e.g., Erb, Shapley, et al., 2006; Maiolino, Nagao, et al., 2008; Steidel, Rudie, et al., 2014; Zahid, Dima, et al., 2014).

Despite the well-established gas-phase MZR, our understanding of the amount of metals that stars incorporate from the gas is less secure. The gas-phase metallicity only indicates the amount of metals in the gas during the time of observation. The metallicity in stars indicates the metal content in stars at their formation. Therefore, measuring stellar metallicity from a galaxy's integrated stellar light can reveal the "star formation history-averaged" galactic metallicity. It is less susceptible than the gas to instantaneous fluctuations. Obtaining stellar metallicities over a range of galaxy masses, i.e., the stellar MZR, can provide insight on the chemical evolution of galaxies complementary to the gas-phase MZR. For example, Y. Peng et al. (2015) compared the stellar MZR of local star-forming galaxies to that of quenched galaxies to study galaxy quenching mechanisms.

A few suites of cosmological hydrodynamical simulations have made quantitative predictions of the stellar MZR in the past few years. Ma, Hopkins, Kasen, et al. (2016) investigated the evolution of both gas-phase and stellar MZRs over redshift from a limited number of galaxies in the Feedback in Realistic Environment (FIRE) cosmological zoom-in simulations. These simulations suggest that the stellar MZR evolves monotonically, with an increase in stellar metallicity of ~ 0.3 dex at fixed stellar mass from $z = 1$ to $z = 0$. De Rossi et al. (2017) presented stellar MZRs at four redshifts derived from a different suite of cosmological hydrodynamical simulations, the Evolution and Assembly of GaLaxies and their Environments (EAGLE). The derived stellar MZRs came from a larger number of galaxies than those in Ma, Hopkins, Kasen, et al.'s (2016) study but with coarser spatial resolution. De Rossi et al. (2017) predicted that the evolution of the stellar MZR is 0.2 dex from $z = 1$ to $z = 0$ at a stellar mass of $10^{9.5} M_{\odot}$, slightly smaller than predicted by Ma, Hopkins, Kasen, et al. (2016).

The classical approach for measuring ages and metallicities of stellar populations is to use spectrophotometric indices such as Lick indices (Faber, 1973; Worthey et al., 1994), where the equivalent widths of some spectral features expected to correlate with metal abundance or age are measured and compared to model predictions. Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) conducted one of the pioneering works using spectrophotometric indices to measure the stellar

MZR from local $z \sim 0$ SDSS star-forming and quiescent galaxies. Gallazzi, Bell, et al. (2014) extended their work to 77 galaxies at $z \sim 0.7$. They found an offset of the stellar MZR by 0.12 dex from $z = 0.7$ to $z = 0$ among the star-forming population but no significant evolution among the quiescent population.

However, the stellar metallicities of star-forming galaxies should be taken with caution since they tend to be more difficult to measure. Large uncertainties and biases could arise from emission line subtractions, a lack of young stars in stellar libraries, and the fact that the most luminous stars are not necessarily the majority of the mass. The uncertainties of metallicities in star-forming galaxies measured by Gallazzi, Charlot, Brinchmann, and S. D. M. White (2006) are generally twice the uncertainties of the metallicities in quiescent galaxies, with a median of $\delta \log Z \sim 0.16$ for star forming galaxies and 0.08 for quiescent galaxies. Moreover, using a similar sample of SDSS galaxies, Panter et al. (2008) found that star-forming galaxies have higher stellar metallicity than the whole sample, while Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) and Y. Peng et al. (2015) found the opposite. Adding further confusion, when metallicities were measured from equivalent widths of UV absorption lines, Sommariva et al. (2012) found that the stellar MZR of star-forming galaxies at $z = 3$ is consistent with that of local galaxies measured by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005), i.e., no significant evolution of the stellar MZR among the star-forming galaxies from $z = 3$ to $z = 0$. This contradicts the conclusion of Gallazzi, Bell, et al. (2014). Nonetheless, the metallicities measured by Sommariva et al. (2012)—using light mainly produced by O stars—might trace different populations from those measured by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005)—using the light from stars of earlier spectral type.

If we focus on the quiescent galaxies whose ages and metallicities can be measured more reliably, a number of recent works have been employing an alternative approach—a full spectrum fitting technique—to determine the ages and metallicities of these quiescent stellar populations. The modeling of full optical–NIR spectra of stellar systems has been advanced in the past decade (e.g., Cid Fernandes et al., 2005; Ocvirk et al., 2006; Walcher et al., 2009; Conroy, Gunn, and M. White, 2009; Conroy and van Dokkum, 2012; Conroy, Graves, et al., 2014). This method is preferred over the use of spectrophotometric indices because it utilizes nearly all of the information from the collected light, resulting in smaller uncertainties.

Though the approach has been used to measure ages and metallicities of both local

and high redshift galaxies, the measurements of higher redshift galaxies are still mainly limited to stacks of spectra due to the lack of sufficient signal in individual spectra. Choi, Conroy, et al. (2014) measured the stellar MZR from stacked spectra of quiescent galaxies ranging from $z = 0$ to $z = 0.7$. The stellar MZRs show possible evidence for evolution with redshift. The MZR measured from stacked spectra can only reflect the median metallicities of the population. It cannot reveal the the scatter of age or metallicity within the population. In fact, Choi, Conroy, et al. (2014) measured very different ages and metallicities of two individual galaxies at $z = 0.8$. The ages and metallicities of the individual galaxies were significantly different from the results from the stacked spectrum at the same redshift. Although the two galaxies were brightest cluster galaxies, the results suggest a possible large scatter within the population. Moreover, measuring galaxy properties in individual galaxies can reveal potentially important correlations between galaxy parameters such as Δ[Fe/H] and age or Δ[Fe/H] and [α/Fe], where Δ[Fe/H] is the deviation in an individual galaxy's metallicity from the MZR of the whole population.

Ultimately, the observations so far reveal no strong evidence for the evolution of the stellar MZR, though the simulations suggest otherwise. If evolution is present, conceivably it has not been detected because earlier observations have been limited to the most massive galaxies. De Rossi et al. (2017) predicted that the evolution of the stellar MZR is 0.2 dex from $z = 1$ to $z = 0$ for a mass of $10^{9.5} M_{\odot}$ but less than 0.05 dex at $10^{10.5} M_{\odot}$ over the same redshift span. Choi, Conroy, et al.'s (2014) sample at $z > 0.4$ is limited to massive galaxies ($M_* > 10^{10.8} M_{\odot}$). The passive galaxies at $z = 0.7$ in Gallazzi, Bell, et al.'s (2014) sample are also massive ($M_* > 10^{10.5} M_{\odot}$) and have large uncertainties in [Fe/H] on the order of ~ 0.2 dex. Onodera, Carollo, Renzini, et al. (2015) used Lick indices to measure the age and stellar metallicity of a stacked spectrum of passive galaxies at $z = 1.6$, the highest redshift at which stellar metallicity has been measured. The masses were also limited to $M_* > 10^{11} M_{\odot}$, where minimal evolution is expected. Since all of the observations at higher redshifts so far are limited to the massive end, it is not surprising that there is no statistically significant evidence for the evolution of the stellar MZR. In fact, the MZR of Gallazzi, Bell, et al. (2014) tentatively suggests a stronger chemical evolution in lower mass galaxies even among the quiescent populations.

This paper is the first in a series to present individual stellar metallicity measurements from $z \sim 0$ to $z \sim 1$. For the first time, we report a detection of $> 5\sigma$ in the evolution of the stellar MZR with redshift based on the stellar MZR of 62 early-type galaxies

from the galaxy cluster Cl0024+1654 at $z \sim 0.4$. To our knowledge, this is the first attempt to measure metallicities using full spectrum synthesis from individual spectra that extend to stellar masses as low as $10^{9.7} M_{\odot}$ beyond the local universe. The major advances in this paper that made the detection possible are (a) the measurement of ages and metallicities in individual galaxies at $z > 0$; (b) the extension of the mass range to low-mass galaxies; and (c) the use of full spectrum synthesis in deriving the parameters. Aside from the evolution with observed redshift (at 0.037 ± 0.007 dex per Gyr), we also detect an even stronger evolution of the stellar MZR with the redshifts at which the galaxies formed (at 0.055 ± 0.006 dex per Gyr).

In Section 4.2, we describe the data and their completeness. In Section 4.3, we describe the method we used to measure ages and metallicities including how it performed as a function of signal-to-noise ratio and when the assumption of the single stellar population was dropped. Further comparisons between our measurements and the measurements from the literature can be found in Appendix 4.A. In Section 4.4, we present the MZR derived from a subsample of local SDSS quiescent galaxies. In Section 4.5, we discuss our results and demonstrate that there is evolution in the stellar MZR with both observed redshift and with formation redshift, i.e., when galaxy ages are taken into account. We also examine the relatively gentle slope found in the observed MZR and how it relates to the feedback strength of galaxies in the observed mass range. Lastly, we discuss the impact of cluster environments on our results in Section 4.6. Throughout this paper, we assume cosmological parameters $\Omega_{\Lambda} = 0.7$, $\Omega_m = 0.3$ and $H_0 = 70$ km s^{-1}Mpc^{-1}.

4.2 Data

We leverage a large imaging and spectroscopic survey of the $z \simeq 0.4$ galaxy cluster Cl0024+1654 (Treu et al., 2003; Moran, R. S. Ellis, Treu, Smail, et al., 2005; Moran, R. S. Ellis, Treu, Salim, et al., 2006; Moran, Miller, et al., 2007; Moran, R. S. Ellis, Treu, Smith, et al., 2007). The study provides comprehensive Keck spectroscopy and Hubble Space Telescope (HST) imaging of Cl0024+17 members. Studying galaxy clusters offers the advantage of being able to obtain a large number of spectra in a few Keck pointings.

Moran, Miller, et al. (2007) provide complete details of the survey. In summary, the HST imaging of Cl0024+17 consists of 39 sparsely sampled WFPC2 images taken in the F814W filter at an exposure time of 4-4.4 ks each. The imaged field spans up to > 5 Mpc from the cluster center, a significantly larger radius

than the cluster's virial radius of 1.7 Mpc. Supplementing the HST imaging, the study also provides infrared imaging in the K_s-band and J-band from the WIRC camera on the Hale 200-inch telescope, optical imaging in the $BVRI$ bands from the 3.6 m Canada-France-Hawaii Telescope, and both near- (NUV) and far-ultraviolet (FUV) imaging on the Galaxy Evolution Explorer (GALEX) satellite. We used the photometric catalog from the survey's Web site[1] which includes all optical and infrared photometric measurements, HST morphologies, and photometric/spectral redshifts. We downloaded the UV catalog separately through the Mikulski Archive for Space Telescopes (MAST). All UV galaxy images were visually examined with the optical counterparts to ensure the correct associations.

The survey used DEIMOS (Faber, Phillips, et al., 2003) on the Keck II telescope to obtain deep spectra of 300 member galaxies to $M_V = -18$. The observations took place between 2003 and 2005. The targets were selected with priority given to known cluster members up to $I = 22.5$ with classified HST morphologies. A total of 16 masks were observed with integration times of 2-4 hrs each. Twelve masks were observed with the 900 line mm^{-1} grating from 2003 to 2004 while the rest were observed with the 600 line mm^{-1} grating in 2005 (Moran, Miller, et al., 2007). All slitlets were 1″ wide, yielding spectral resolutions of $R \sim 2000 - 3000$. All masks were centered at $\sim$ 6200 Å, providing a spectral range of 3500 to 5500 Å in the rest frame. The spectroscopic sample is > 65% complete for objects with $m_{\mathrm{F814W}} < 21.1$.

The DEIMOS spectra were reduced using the spec2d DEIMOS data reduction pipeline (Cooper, J. A. Newman, et al., 2012; S. F. Newman, Genzel, Förster Schreiber, et al., 2013) adapted by Kirby, Simon, et al. (2015). Each spectrum was flat-fielded, wavelength-calibrated, sky-subtracted and telluric-corrected.

We selected subsamples of quiescent galaxies with an average signal-to-noise ratio (S/N) greater than 8 Å^{-1} in the observed frame to ensure that we can break the degeneracy between age and metallicity (see Section 4.3). We define quiescent galaxies as those with rest-frame equivalent widths (EWs) of [O II] λ3727 smaller than 5 Å and rest-frame FUV−V colors larger than 3. The EW limit is roughly equivalent to having a cut in specific star formation rate (sSFR)[2] at approximately $10^{-11} M_\odot$ yr^{-1}. The color cut is intended to further minimize contamination by star-forming

[1] `http://www.astro.caltech.edu/clusters/`

[2] The sSFR approximation is based on the SFR([O II]) calibration from Kewley, Geller, et al. (2004), the stellar mass-to-light (M/L_B) ratios from Bell et al. (2001) and the mean rest-frame (U-B) color $\sim$ 1.5 of the sample.

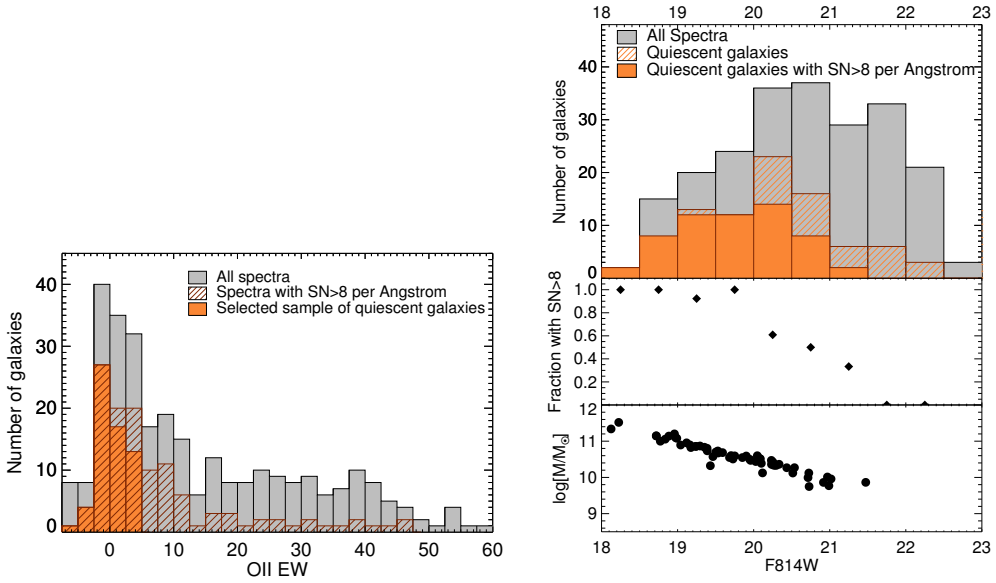

Figure 4.1: Characteristics of the observed spectra in the Cl0024+17 cluster. (*Left*) Histograms of [O II] λ3727 rest-frame equivalent widths. In this study, we use samples of quiescent galaxies (EW $<$ 5 Å and FUV$-$V $>$ 3) with S/N $>$ 8 Å^{-1}. (*Right*) Histograms of galaxies as a function of F814W magnitude. The bottom panel shows magnitude as a function of stellar mass. Our sample is $\sim$ 50% complete at $M_* > 10^{9.7} M_\odot$.

galaxies. Specifically, the cut eliminates galaxies with star formation more recent than $10^7 - 10^8$ yr (see Moran, R. S. Ellis, Treu, Salim, et al., 2006). The spectral S/N is estimated as the inverse of the average ratio between the absolute deviation of the observed spectrum from the best-fit spectrum of all pixels in continuum region. We quote the S/N per Å rather than per pixel.

The [O II] EW, FUV$-$V color, and S/N criteria reduced the sample from the original total of 300 observed DEIMOS spectra to 62 quiescent spectra with sufficient S/N. Figure 4.1 shows the fraction of the selected sample to the parent sample as a function of F814W magnitude and stellar mass. Stellar masses and rest-frame colors were derived from available photometry using the SDSS KCORRECT software version v4_3 (Blanton et al., 2007), which assumes Bruzual et al. (2003) population synthesis models and the Chabrier (2003) stellar initial mass function. The fraction of quiescent galaxies with S/N $>$ 8 Å^{-1} is $\sim$ 75% for F814W$<$ 21.1. Since the parent survey is $>$ 65% complete at the same magnitude range for the DEIMOS spectra, our final sample is therefore $\sim$ 50% complete for F814W$<$ 21.1 or $M_* \gtrsim 10^{9.7} M_\odot$.

4.3 Model Fitting

In this section, we describe the full-spectrum fitting technique used to derive ages and metallicities. Section 4.3 describes the stellar population synthesis models used in this work. In Section 4.3, we detail our fitting technique. We test the accuracy of our measurements against our assumptions in Section 4.3.

Model Spectra

We derived stellar metallicities and ages of the galaxies using the full-spectrum fitting technique via stellar population synthesis (SPS). In the past decade, SPS models have been refined and often used as alternative tools to photometric indices to derive stellar population parameters such as stellar mass, star formation history, metallicity, and age (e.g., Schiavon et al., 2006; Walcher et al., 2009; Choi, Conroy, et al., 2014; Fumagalli et al., 2016). The commonly used SPS models include Bruzual et al. (2003, BC03), Fioc et al. (1999, PEGASE), Maraston (2005, p. M05), and Conroy, Gunn, and M. White (2009, FSPS). The main advantage of using SPS over photometric indices is that it utilizes information from the whole spectrum simultaneously instead of using portions with the strongest stellar absorption features. Consequently, spectra with lower signal-to-noise ratios can be used to achieve the same level of precision.

We adopted the Flexible Stellar Population Synthesis (FSPS) models (Conroy, Gunn, and M. White, 2009) among the publicly available SPS models to derive the stellar population parameters from our sample. FSPS utilizes the most recent model of the Padova stellar evolution tracks (Marigo, Girardi, et al., 2008). The model includes treatments of thermally pulsating asymptotic giant branch (TP-AGBs) and an option to include horizontal-branch (HB) and blue straggler (BS) stars. The TP-AGBs are particularly important because leaving them out can result in systematic differences in age by a factor of 2 (Maraston, 2005; Conroy, Gunn, and M. White, 2009). Lastly, the model offers flexibility in modeling the spectra that match the spectral resolution and the ranges of stellar parameters in our study.

We generated the templates of single stellar population (SSP) spectra using the FSPS code version 3.0. The SSP spectra were generated with the Kroupa IMF (Kroupa, 2001) and the MILES spectral library (Sánchez-Blázquez et al., 2006). We chose the MILES spectral library because it spans a wide range in the stellar parameters, i.e., $\log g$ and [Fe/H], that are suitable for measuring galaxies at higher redshifts with lower metallicity. In addition, the library has a spectral resolution of

2.3 Å FWHM across the wavelength range of 3525 to 7500 Å, comparable to that of our spectra. The spectra were interpolated from 22 modeled metallicities ranging from $\log Z = -1.98$ to 0.2, corresponding to the metallicity values of the Padova isochrones. The age ranges from 0.3 Myr to 14 Gyr. The rest of the parameters were set to the default mode in generating SSP spectra, which means dust, blue HB, and BS stars were excluded. Because dust absorption mainly affects the continuum of the spectra but not the absorption lines, and because we remove the continuum from the observed spectra in our fitting procedure, we omitted dust for simplicity. The omission of blue HBs and BS stars should not affect the derived ages and metallicities because these stars mainly contribute to ultraviolet wavelengths, which are not observed in our spectra.

Measurements of metallicities and ages

We iteratively fit each spectrum to FSPS models via χ^2-minimization. First, we created a mask for continuum normalization where all emission lines, strong absorption lines, and telluric regions are masked out. We continuum-normalized each galaxy spectrum with B-spline fitting with breakpoints at every $\sim$ 100 Å. We then created a different mask for the fitting procedure. In this mask, the Mg b triplet and emission lines (if [O II] was detected in emission) were masked out from the continuum-normalized spectra. The Mg b triplet was masked out so that the measured [Fe/H] would better reflect iron abundance rather than magnesium abundance. For each iteration, we simultaneously fit for four parameters: [Fe/H], age, velocity dispersion, and redshift. The priors were uniform for [Fe/H] and age. [Fe/H] was in the range of $[-1.98, 0.2]$, while the age was in the range of 0.3 Myr to the age of the universe at the galaxy's observed redshift. We set the prior for velocity dispersion according to the Faber-Jackson relation (Faber and Jackson, 1976) between the velocity dispersion and the stellar mass measured by Dutton, Conroy, et al. (2011), with a range of ± 0.4 dex, which is large enough to capture any uncertainties and evolution with redshift (Dutton, van den Bosch, et al., 2011).

The spectral fitting proceeded as follows. First, we used the IDL code MPFIT (Markwardt, 2012) to fit the continuum-normalized observed spectrum with the SSP spectra. In this first fitting iteration, all model SSP spectra were continuum-normalized in the same manner as the observed spectrum using the same continuum mask. By fitting continuum-normalized spectra in the first iteration, we can bypass uncertainties from flux calibration, dust absorption, and uncertainties of the continuum flux in the modeled SSPs. This is also important because the MILES library

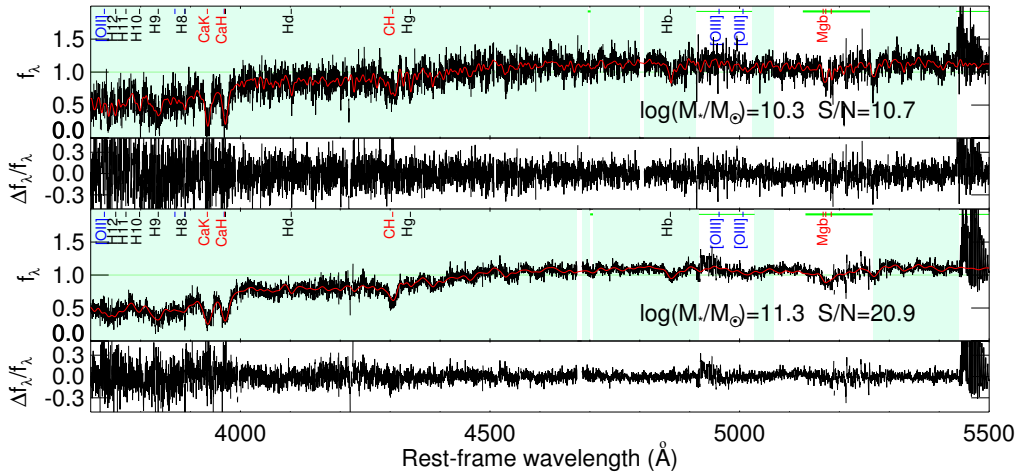

Figure 4.2: Examples of observed $z \sim 0.4$ spectra (black) and the corresponding best-fit models (red). The flux is continuum-normalized with an applied "synthesized" continuum. The model spectra are normalized by their median flux and smoothed to the instrumental resolution and best-fit velocity dispersion. The teal background shows the spectral regions used for spectrum modeling while the white background show the spectral regions that are masked out. The green bars show regions with strong telluric absorption lines. We display the fractional residuals in the bottom panel of each spectrum. The measured metallicities and ages of the two galaxies are [Fe/H] = $-0.14^{+0.13}_{-0.09}$, $-0.15^{+0.04}_{-0.06}$ and Age = $3.2^{+1.1}_{-0.6}$, $6.3^{+0.7}_{-0.8}$ Gyr, respectively. The uncertainties include the systematic uncertainties from the age–metallicity degeneracy.

provides spectra that have been normalized to unity.

For the rest of the iterations, we did not continuum-normalize any model spectra in order to minimize any alteration to the model spectra. Instead, we applied a "synthesized" continuum curve to the observed continuum-normalized spectrum. To do so, (1) we divided the observed, continuum-normalized spectrum by the best-fit SSP model spectrum from the previous iteration; (2) we fit the result from (1) with a B-spline using the same continuum mask; (3) we divided the observed continuum-normalized spectrum by the continuum curve from (2) to create an observed spectrum with a "synthesized" continuum; (4) we re-fit the resulting spectrum with the SSP model spectra—including the continuum shape—using the MPFIT code. The process (1)-(4) was repeated for one hundred iterations, more than enough for the parameters to converge. We show examples of the observed spectra and their best-fit spectra in Figure 4.2.

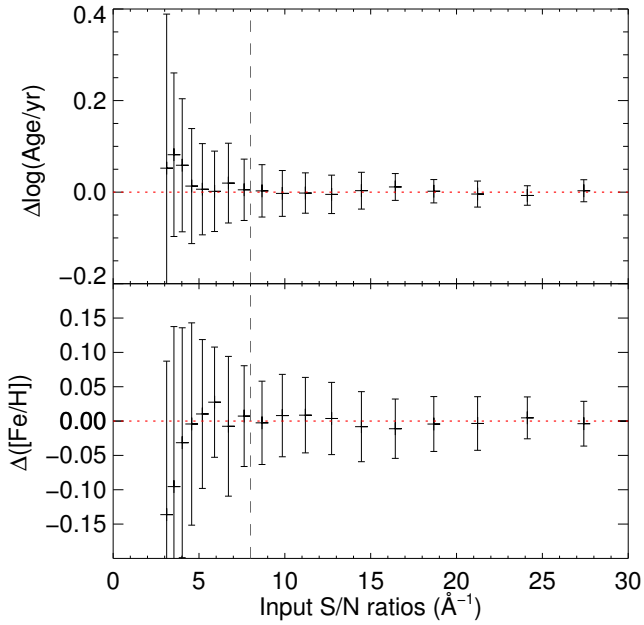

Figure 4.3: Accuracy in measurements of ages and metallicities as a function of S/N when assuming a SSP. Each data point shows the mean and standard deviation of 20 mock SSP spectra with a certain S/N. At S/N $\gtrsim$ 8 Å^{-1}, we can recover the metallicities within 0.05 dex.

Accuracy of metallicities and ages

In this section, we first examine the statistical uncertainties as a function of signal-to-noise ratio of the spectra and when the assumption of SSP is dropped. We then explore the systematic uncertainties that arise from the age–metallicity degeneracy. Lastly, we refer the reader to Appendix 4.A for comparisons between our age and metallicity measurements and those in the literature.

Dependence on signal-to-noise ratios

To investigate how the observed signal-to-noise ratios (S/Ns) influence the uncertainties in the measured ages and metallicities, we tested our spectral fitting code on a set of mock DEIMOS spectra with different S/Ns. We first adopted the SSP assumption. We created a mock SSP spectrum from the FSPS code with $\log Z = -0.1$ dex and age of 3 Gyr. These numbers were chosen to be representative for our $z \sim 0.4$ data. The spectrum was smoothed to have a velocity dispersion of FWHM = 250 km/s and the same spectral range and resolution as a typical DEIMOS spectrum. Gaussian noise was added to the spectrum to create 20 spectra for each S/N ranging

from $\sim$ 3 to 30 Å^{-1}. We then multiplied the spectra with a telluric transmission curve and the DEIMOS instrumental throughput to mimic the observed spectra.

We found that given the SSP assumption, we can measure [Fe/H] and age well to $\sim$ 0.05 dex precision when the S/Ns of the spectra are $>$ 8 Å^{-1}, which is the minimum S/N in our sample. The uncertainties in [Fe/H] and age as a function of S/N are shown in Figure 4.3. The fluctuation in the measurements decreases rapidly as a function of S/N. However, once the S/N is higher than $\sim$ 8 Å^{-1}, the fluctuations decrease slowly. The level of the fluctuation does not change much when we change the age of the mock spectra from 3 Gyr to 8 Gyr. With uncertainties $\lesssim$ 0.05 dex, we conclude that we only minimally suffer from statistical uncertainties.

Validity of the SSP assumption

Next, we test the validity of the SSP assumption. We created a set of mock spectra of a composite stellar population (CSP) with an exponentially declining star formation history $\psi(t') \propto exp(-t'/\tau)$ with the following parameters: τ = 1 Gyr, $\log(Z/Z_\odot) = -0.2$, and an elapsed time ranging from 1 to 8 Gyr since the onset of the star formation. We added Gaussian noise so that the S/N is 12 Å^{-1}, the average S/N of our spectra. We applied telluric features, the instrumental throughput, and a smoothing kernel to the mock spectra in the same manner as in Section 4.3. We note that in these spectra, all stars have the same metallicity and should be interpreted as the population's light-weighted metallicity.

We then fit the CSP model spectra with SSP models. In general, we recovered the CSP age and metallicity within $\sim$ 0.1 dex precision when most of the star formation has been quenched. The measured metallicities and ages are shown in Figure 4.4. The gray curve in the top figure shows the shape of the exponentially declining star formation rate. The red dashed line is plotted as a guide for the "light-weighted" age of the population at elapsed time t via

$$\mathrm{Age}(t) \sim \frac{\int_0^t (t - t')L(t')\psi(t')dt'}{\int_0^t L(t')\psi(t')dt'}$$

where $L(t')$ is the integrated light in the wavelength range of rest-frame 3700 to 5500 Å produced by SSP stars of age t'.

Although the integrated light $L(t')$ accounts for the massive stars that died before reaching age t', the integration is over every star that has formed ($\psi(t')dt'$) and does not account for the stars that have died. Therefore we expect this "light-weighted" age

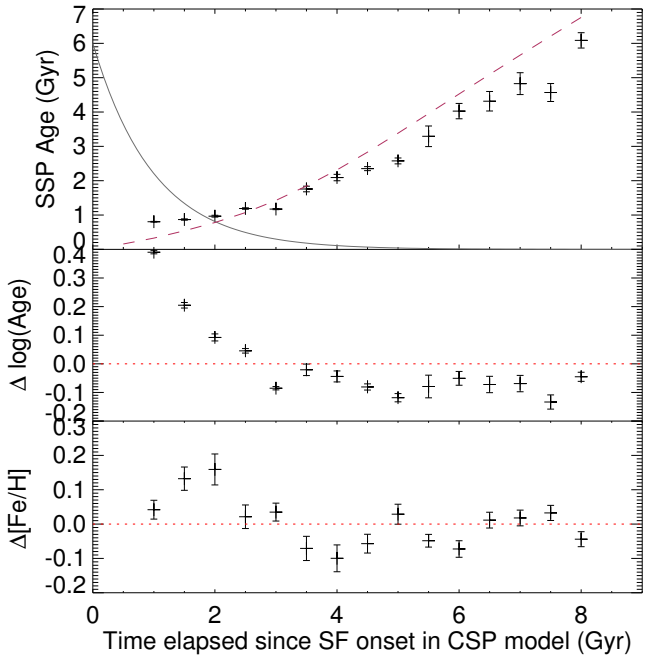

Figure 4.4: Accuracy in measurements of ages and metallicities when SSP is not assumed. (*Top*) Measured SSP ages from the mock spectra of composite stellar populations. The spectra have S/N = 12 Å^{-1} and the same star-formation history but were observed at different amounts of time elapsed after the onset of the star formation. The solid curve shows the assumed exponentially declining star formation history. The red dashed line shows the light-weighted age calculated over the 3700-5500 Å range. (*Middle and bottom*) The deviation of the measured ages and metallicities from the input values. The middle panel shows the difference between the measured SSP ages and the red dashed line in the top panel.

to overestimate the time elapsed since the beginning of star formation, especially for older populations. The age overestimate in old populations might partially explain the behavior shown in the middle plot. The SSP-equivalent ages that we measured are consistently younger than their respective "light-weighted" ages of the CSP population.

The measured metallicities fluctuate well within 0.1 dex of the true answer, and they do not seem to be affected by assuming an SSP rather than a CSP. We conclude that we can measure metallicity to a precision of $\sim$ 0.1 dex under the SSP assumption. On the other hand, we likely underestimated ages by an amount less than 0.1 dex as long as the majority of the star formation has been quenched, i.e., the light-weighted age is greater than $\sim$ 1.5 Gyr. Because our sample is comprised of quiescent galaxies, we expect that our age and metallicity measurements are not greatly affected by the

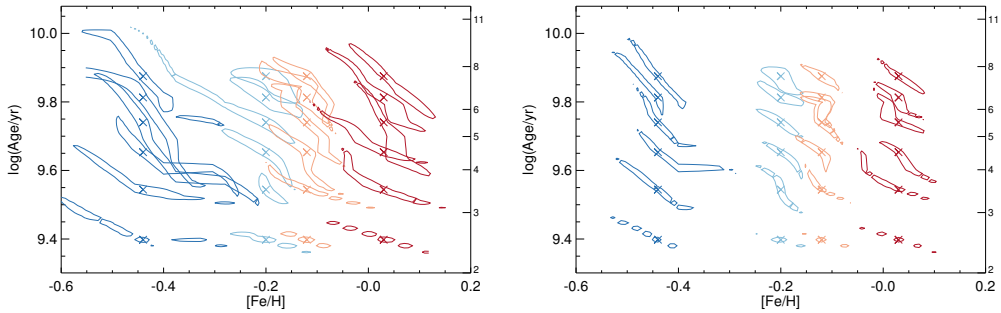

Figure 4.5: Uncertainties in the measured ages and metallicites according to the age–metallicity degeneracy. Each contour shows the 1σ range in age and metallicity. The true ages and metallicities are shown as cross marks. The S/N of each spectrum is 10 Å^{-1} (left) and 25 Å^{-1} (right).

SSP assumption.

Our results regarding the SSP assumption are similar to the results by Trager et al. (2009). The authors concluded that an SSP-equivalent metallicity is an excellent tracer of the light- or mass-weighted metallicities, whereas an SSP-equivalent age generally biases toward values younger than the true mass- and light-weighted ages. The level of bias in ages at the level of ~ -0.1 dex compared to the true light-weighted age is also similar to ours. Trager et al. argued that hot young stars contribute minimally to the metal lines but heavily to the Balmer lines relative to old stars.

Age–metallicity degeneracy

For each galaxy, we estimated the systematic uncertainty arising from the age–metallicity degeneracy. As shown in Figure 4.4, the statistical uncertainties obtained from MPFIT underestimate the level of total uncertainty. To estimate total uncertainty, we created a mock SSP spectrum with the same age and metallicity as each observed galaxy. Gaussian noise was added to the spectrum to reach the same S/N as the observed spectrum. We then compared the noised spectrum to a 100×120 grid of noise-less SSP spectra with a range [0.5,13] Gyr in age and [-0.8,0.2] dex in [Fe/H] and calculated a χ^2 array for the noised mock spectrum. All SSP spectra were smoothed to achieve a velocity dispersion of 250 km/s FWHM convolving with the typical resolution of an SDSS or a DEIMOS spectra.

The uncertainties in age and metallicity are calculated by marginalizing the 2–D posterior probability distribution obtained from the χ^2 array. The uncertainties for

upper (lower) limits are the differences between the values at 84th (16th) percentile and the 50th percentile in the posterior probability distributions. Since we calculate these uncertainties specifically to the S/N of each galaxy, we take these uncertainties as the total uncertainty of each measurement. The uncertainties quoted in the subsequent text and figures refer to these systematic uncertainties.

In general, the uncertainties in ages and metallicities are neither Gaussian nor symmetric. Figure 4.5 shows contours of 1σ uncertainty based on the 2–D posterior probability distribution functions of the age–metallicity degeneracy at S/N= 10 and S/N= 25 Å^{-1}. The average total uncertainties for upper and lower limits are +0.11 and −0.14 dex for [Fe/H] and +0.12 and −0.11 dex for age, respectively. We list all our measurements of ages and metallicities in Table 4.1.

4.4 Mass–Metallicity Relation of Local Galaxies

In this section, we report the MZR of SDSS local galaxies measured with our method. We will use this MZR as a reference to compare with the MZR of $z \sim 0.4$ galaxies in Section 4.5.

Figure 4.6 shows the stellar mass–stellar metallicity relation we measured for local quiescent SDSS galaxies. We selected a subsample of 155 quiescent galaxies from Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005). The selection criteria and detailed comparisons to the age and metallicity measurements of Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti are described in Appendix 4.A. In order to ensure that the stellar masses of both the $z \sim 0$ and $z \sim 0.4$ samples are on the same scale, we remeasured the stellar masses of the local galaxies with the KCORRECT code using the SDSS u, g, r, i, z photometry. The MZR of local galaxies shows the expected relation; metallicity increases with stellar mass. However, the "knee," or where the slope in the MZR changes, is less visible from our measurements than from Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005). The MZR is generally consistent with the MZR of the complete galaxy population (both star-forming and quiescent galaxies) analyzed by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti at stellar mass above $M_* \approx 10^{10.3} M_\odot$. However, at lower stellar masses, the MZR of the subsample of quiescent galaxies shows metallicities higher than those of Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti.

We consider three functions to approximate the MZR. First, we try fitting with the three-parameter logarithmic function proposed by Moustakas et al. (2011) to

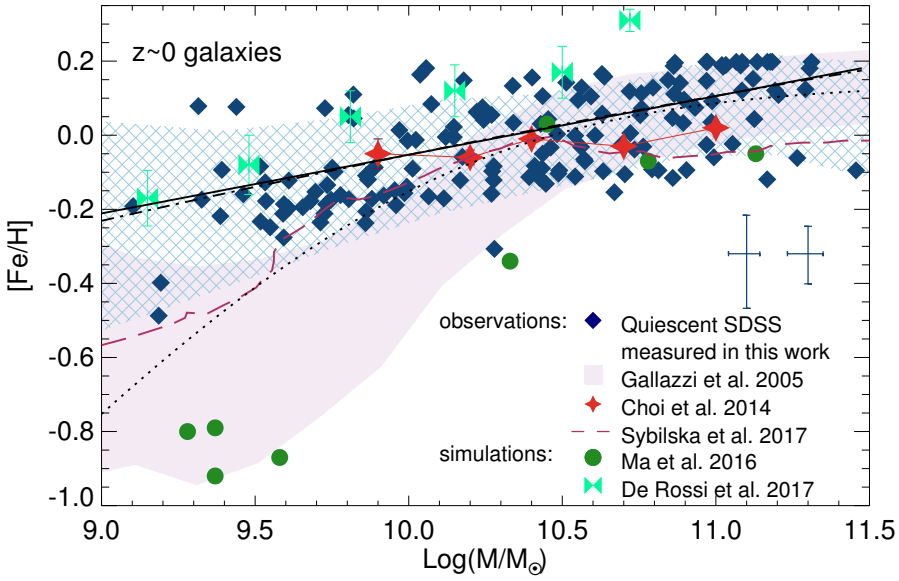

Figure 4.6: The stellar mass–metallicity relation of local quiescent galaxies. Each navy-blue diamond shows the measured metallicities in this work. The light-blue hatched strip shows the average and standard deviation of metallicities in each mass bin. The two error bars to the right of the plot show median uncertainties for galaxies with stellar mass lower (left) and higher (right) than $10^{10} M_\odot$. The black solid curve is the best-fit linear function while the black dash-dot and dotted lines and are the best-fit quadratic and logarithmic (Equation 4.1) function, respectively. The light-purple solid strip shows the average relation from Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005). The plot also shows local MZRs measured in other works (Choi, Conroy, et al., 2014; Sybilska, Lisker, et al., 2017) and the local MZRs found in the FIRE and EAGLE hydrodynamical simulations (Ma, Hopkins, Kasen, et al., 2016; De Rossi et al., 2017).

describe the gas-phase MZR:

$$[\mathrm{Fe/H}] = [\mathrm{Fe/H}]_0 - \log[1 + M_*/M_0]^\gamma \tag{4.1}$$

We then fit with a quadratic function, another proposed form of gas-phase MZR (e.g. Tremonti et al., 2004), and a linear function. For each function, we fit with 1000 iterations of the Monte Carlo random sampling method because the uncertainties of our measurements are not Gaussian. In each iteration, we resample according to the probability distribution in [Fe/H] of each galaxy and find the best-fit function by minimizing the chi-square error statistic.

The MZR of the subsample of quiescent galaxies is best described by a linear fit with a modest slope. The best-fit logarithmic function (the bottom-most dash-dot

line in Figure 4.6) performs the worst in terms of χ^2, at approximately 1.6 times the minimum χ^2 of the two other functions. Both quadratic and linear functions fare equally well based on the minimum χ^2. However, according to the Akaike information criterion (Burnham et al., 2003), the linear function is a better choice because it is less complex and minimizes the loss of information. The slope of the best linear fit is $\sim 0.16 \pm 0.02$ dex in [Fe/H] per log mass—consistent with the slope found by Gallazzi, Charlot, Brinchmann, and S. D. M. White (2006). The slope indicates that an increase in galaxy mass by a factor of 10 corresponds to an increase in metallicity by a factor of 1.4.

When compared with the MZR measured by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005), higher metallicities in the low-mass galaxies can directly result from the systematic difference in the metallicity measurements. We tried plotting the MZR using the stellar masses reported in Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005), which were measured from spectral indices and z-band photometry as compared to the KCORRECT code used in this paper. Because the mass measurements from both methods are generally consistent within 0.1 dex, the best-fit linear function does not depend on the difference in the mass measurement methods. The systematic difference in metallicity measurement is therefore the cause of the difference in the MZR at the lower mass end.

As discussed in Appendix 4.A, for the same galaxies with low metallicities, we measured their [Fe/H] to be higher than what Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) measured. Since galaxies with low metallicities are mainly less massive galaxies, this results in the MZR lying above that shown by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005). The MZR with higher metallicities at the lower-mass end found here is consistent with the MZRs found in stacks of local spectra in Choi, Conroy, et al. (2014) and in individual local early type galaxies via IFU observations by Sybilska, Lisker, et al. (2017). The metallicities measured by Sybilska, Lisker, et al. (2017) were measured via spectroscopic indices using the SSP models from Vazdekis, Coelho, et al. (2015), which is also based on the MILES library. Because the measurements in Choi, Conroy, et al. (2014), Sybilska, Lisker, et al. (2017), and this work utilize the same stellar library, it is very likely that the choice of stellar library used caused the systematic differences in the metallicity measurements compared to Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005).

Whether the higher metallicity at the lower-mass end of the MZR can be additionally

caused by differences in SFH between the two samples, on top of the measurement methods, is still ambiguous. There is observational evidence that the stellar MZR of early type galaxies might differ from that of late-type galaxies. Although categorizing galaxies based on their morphologies is not necessarily the same as categorizing based on their SFH, the two properties closely correlate with each other in both local and high-redshift galaxies (e.g., Kriek et al., 2009; Wuyts et al., 2011; Lee et al., 2013). Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) found that morphology, described by a concentration index, is responsible for a difference as large as $\sim$ 0.6 dex in [Fe/H] in low-mass galaxies. The difference between the two populations in more massive galaxies is minimal.

On the other hand, theoretical work suggests that there should not be a significant difference between star-forming and passive galaxies. Okamoto et al. (2017) created separate MZRs for passive and star-forming galaxies based on the Illustris simulation (D. Nelson et al., 2015, Illustris-1) and the EAGLE simulation (McAlpine et al., 2016). Okamoto et al. (2017) showed that the difference between passive and star-forming galaxies is $\sim$ 0.05 dex or smaller at any fixed mass. Furthermore, they found that the shape of the MZR does not depend on galaxy type.

Regardless of the true shape of the MZR at smaller masses, we will use our measurements in Figure 4.6 as a point of reference when comparing with those of higher redshift. Comparing to our own measurement of the local galaxy populations reduces the systematic uncertainties introduced from measurements and sample selection.

4.5 Evolution in the mass–metallicity relation

In this section we show the main result of this work. In Section 4.5, we report the MZR of individual $z \sim 0.4$ galaxies in which we, for the first time, detect an evolution in the stellar MZR with observed redshift. In Section 4.5, we explain the scatter in the MZR and report an even greater evolution of the MZR when the age of galaxies is taken into account. Lastly, in Section 4.5, we discuss the meaning of the MZR slope and what it reveals about feedback in galaxies.

Evolution of the MZR with observed redshift

We plot the stellar MZR of the $z \sim 0.4$ sample in Figure 4.7. The $z \sim 0.4$ MZR matches well with that measured by Choi, Conroy, et al. (2014) at the same redshift at $M_* > 10^{10.5} M_\odot$. However, in contrast to Choi, Conroy, et al. (2014), we measured the MZR of individual (rather than stacked) galaxies and extended the MZR relation

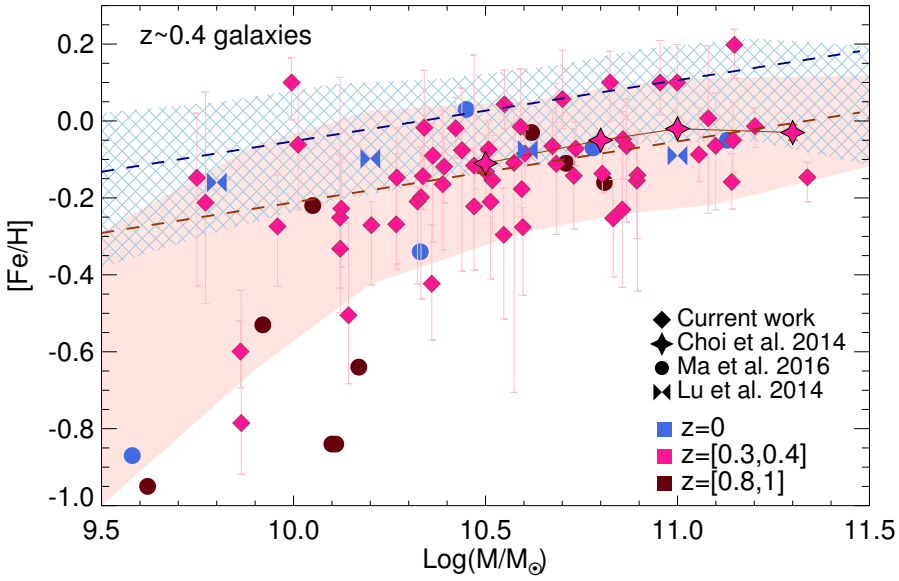

Figure 4.7: Stellar MZR relation of $z \sim 0.4$ quiescent galaxies. The solid orange and blue hatched strips show the average and standard deviation of metallicities in each mass bin of $z \sim 0.4$ and local galaxies (same as Figure 4.6). The upper dashed line and lower dashed line show the best-fit linear functions, where slopes were fixed to the common value, to the local and $z \sim 0.4$ galaxies, respectively. The best-fit parameters are shown in Equation 4.2 and 4.3. We also show the predicted MZR from the FIRE simulations (Ma, Hopkins, Kasen, et al., 2016) and from semi-analytic models with a constant mass-loading factor (Lu, Wechsler, et al., 2014). The plot is color coded by galaxy redshift.

almost 1 dex lower in stellar mass, at $10^{9.7} M_{\odot}$.

Based on the best-fit linear relations to the MZRs at $z \sim 0$ and $z \sim 0.4$, we find an evolution of the MZR with observed redshifts at greater than a 5σ significance level. As in Section 4.4, we fit a linear function to the measured $z \sim 0.4$ MZR via the Monte Carlo method. We then use the analysis of covariance to compare the best-fit linear functions of the MZR at $z \sim 0$ and $z \sim 0.4$. First, we check if the slopes of the the two linear functions are different. The best-fit slope for the $z \sim 0.4$ population is 0.15 ± 0.03 dex per log mass. When compared to the slope of 0.16 ± 0.02 dex per log mass of the $z \sim 0$ population, they are the same within $\sim 1\sigma$ significant. Therefore, we conclude that the slopes of the MZRs at two redshifts are not significantly different. We then test for the evolution in the normalization values (the constant terms). To do so, we re-fit linear functions to the two MZRs using a common fixed slope, equal to the weighted-mean slope of 0.16 dex per log mass.

The best-fit linear equations when the slopes are fixed are

$$\langle[\mathrm{Fe/H}]\rangle = (-0.05 \pm 0.01) + 0.16 \log\left(\frac{M_*}{10^{10} M_\odot}\right) \tag{4.2}$$

for the MZR of $z \sim 0$ quiescent galaxies, and

$$\langle[\mathrm{Fe/H}]\rangle = (-0.21 \pm 0.02) + 0.16 \log\left(\frac{M_*}{10^{10} M_\odot}\right) \tag{4.3}$$

for the MZR of $z \sim 0.4$ quiescent galaxies. The difference in the constant terms is 0.16 ± 0.03 dex, which is significant at greater than a 5σ level.

The observed metallicity evolution is larger than but consistent with the predictions from hydrodynamical simulations. For now, we focus on the shifts in the MZRs with redshift and ignore the slopes predicted in simulations when comparing to our observations. The observed metallicity evolution of 0.16 ± 0.03 dex from $z \sim 0.4$ to $z \sim 0$ translates to an increase of metallicity at 0.037 ± 0.007 dex per Gyr. Based on the FIRE hydrodynamical simulations, Ma, Hopkins, Kasen, et al. (2016) predicted that the stellar metallicity evolution from $z = 0.4$ to $z = 0$ should be 0.13 dex in [Fe/H]. The observed evolution is slightly larger than but consistent within 1σ with the evolution predicted from the FIRE hydrodynamical simulations. We found a larger discrepancy when comparing the observed evolution to the predicted evolution from De Rossi et al. (2017) base on the EAGLE hydrodynamical simulation. De Rossi et al. predicted evolution of 0.11 ± 0.09 dex from $z = 1$ to $z = 0$ in $10^{10} M_\odot$ galaxies (see their Figure 5), which is $\sim 0.014 \pm 0.012$ dex per Gyr. Our observed evolution is therefore also greater but, due to the large uncertainty in the predicted evolution, is consistent within 2σ.

The observed evolution of stellar MZR likely emphasizes the importance of metal recycling in galaxies. Among the hydrodynamical simulations, the FIRE simulations predicted the strongest evolution of the stellar MZR over redshift (Ma, Hopkins, Kasen, et al., 2016), and is the most consistent with our observations. As explained in Ma, Hopkins, Kasen, et al. (2016), this is because the simulated galaxies in the FIRE simulations are able to retain more metals, resulting in higher increase in metallicity over redshifts. This is particularly true in galaxies with stellar mass above $10^{10} M_\odot$, where the retained fractions of metals in the halos are almost unity. The reason that different simulations achieve different metal retention fractions lies in the physical models on which the simulations are based. Many of the cosmological simulations (including the EAGLE simulations) adopt 'sub-grid' empirical models of galactic winds and stellar feedback, where fractions of gas are forced to

escape the galaxy due to energy injection from supernovae and stellar winds. In contrast, the FIRE simulations adopt a Lagrangian formulation of smooth particle hydrodynamics (Hopkins, Kereš, et al., 2014), where metallicities are derived from tracked individual star particles that can be ejected and, importantly, accreted back to the galaxy. As a result, the metal recycling effect is better captured in the FIRE simulations than those that assume 'sub-grid' models.

However, complex and more realistic simulations such as the FIRE simulations are computationally expensive and still limited in terms of sample size, which might explain the small discrepancy between the observed and predicted amount of evolution. The predicted evolution of 0.13 dex from $z = 0.4$ to $z = 0$ in Ma, Hopkins, Kasen, et al. (2016) came from fitting a linear function to the MZRs over a wide mass range, from $M_* \sim 10^4 M_\odot$ to $M_* \sim 10^{11.5} M_\odot$ at different redshifts. If we limit the mass range to above $10^{9.7} M_\odot$, as in our observations, there are only 4 simulated galaxies at $z = 0$ and 8 simulated galaxies at $z = 0.8$. With this limited mass range and number of sample size, the evolution of the MZR in the simulated galaxies is is 0.2 ± 0.6 dex at $10^{10} M_\odot$ from $z = 0.8$ to $z = 0$ or 0.03 ± 0.09 dex per Gyr. Moreover, the FIRE simulations do not include feedback from possible active galactic nuclei (AGNs), which can potentially affect the metallicities in massive galaxies with $M_* > 10^{11} M_\odot$ (Ma, Hopkins, Kasen, et al., 2016). Although the evolution from $z = 0.8$ to $z = 0$ in the FIRE simulation is not significant, it is consistent with the better constrained values from our observations.

We note that the galaxies in both the EAGLE and the FIRE simulations are not necessarily passive, whereas galaxies in our sample are. The fact that the observed evolution of the MZR with redshift is consistent with the simulations does not have any implication on whether the two populations' metallicities are the same at any given redshift (the latter has been suggested by Okamoto et al. (2017)). In this section, we only compared the magnitude of the change of metallicity with redshift, but not the metallicities themselves. In fact, none of the metallicity values are consistent. Different suites of simulations predict different MZR normalizations at each redshift, none of which are consistent with each other or with our observations (see Figures 4.6 and 4.7).

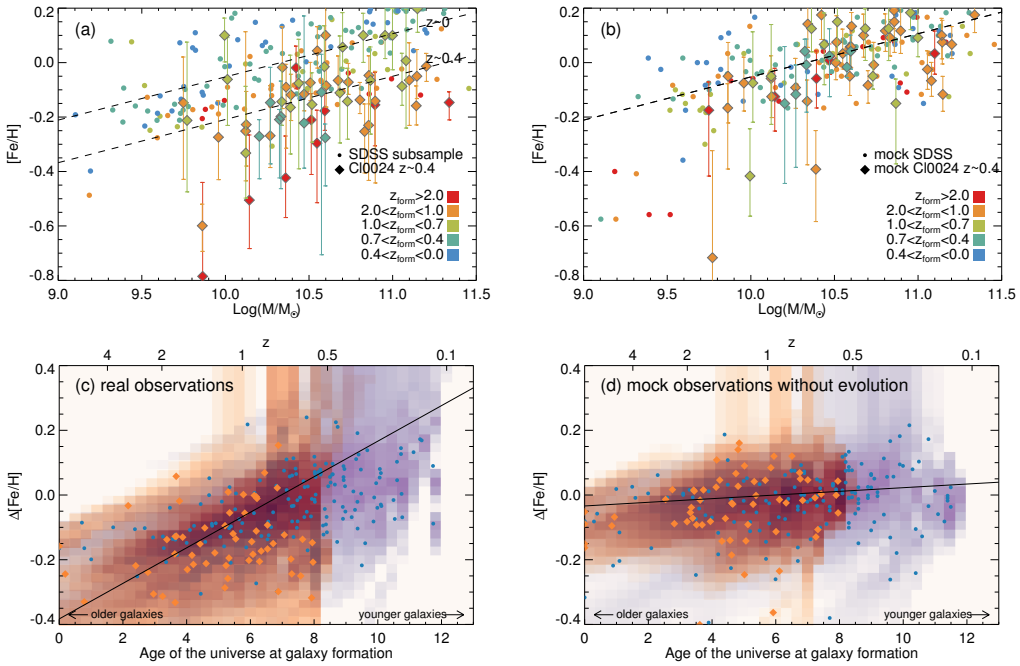

Figure 4.8: (a) Stellar MZRs of both local and $z \sim 0.4$ galaxies color-coded by their SSP-equivalent formation redshifts. The black dashed lines show the best-fit linear functions of the data at $z \sim 0$ and $z \sim 0.4$. (b) Mock observation when the "true" metallicity of each data point is assumed to be a linear function of its mass (the $z \sim 0$ black dashed line). The "observed" ages and metallicities of the mock galaxies were determined by the age–metallicity degeneracy. (c) and (d) The deviation of metallicities from the $z \sim 0$ best-fit linear function in observed data and mock data. The orange diamonds and the blue dots represent data points from $z \sim 0.4$ and $z \sim 0$, respectively. The underlying shaded purple (orange) colors are the co-added probability distribution of individual data points from $z \sim 0$ ($z \sim 0.4$) sample. The slope in the mock data (d) is caused by the age–metallicity degeneracy and is significantly smaller than the slope in the observation (c), which suggests an evolution of the MZR with formation redshift.

Can galaxy formation time explain the evolution of the MZR with observed redshift?

We now measure the intrinsic scatter in the MZR and test for the correlation between the scatter and the galaxy's redshift of formation.

We repeat the linear fit to the stellar MZR with an additional parameter, an intrinsic variability σ_v, by minimizing the negative-log-likelihood

$$L = -\sum_i \log\left[P(\Delta[\mathrm{Fe/H}]_i) * \mathcal{N}(0,\, \sigma_v^2)\right]\Big|_{\Delta[\mathrm{Fe/H}]_i=0}$$

where $P(\Delta[\mathrm{Fe/H}]_i)$ is the probability of the difference between the observed metallicity and the model linear equation $[\mathrm{Fe/H}]_i - (a \log \frac{M_{*,i}}{10^{10} M_\odot} + b)$, which has the same shape as the probability of each observed metallicity $P([\mathrm{Fe/H}]_i)$. This equation means that the probability of the deviation of each observed metallicity from the linear model is equal to the convolution between the probability function inherited from the age–metallicity degeneracy $P[\mathrm{Fe/H}]_i$ and the Gaussian probability of an intrinsic scatter of size σ_v.

The intrinsic scatter in [Fe/H] for local and $z \sim 0.4$ sample are both 0.07±0.01 dex. The sample at both redshifts have consistent intrinsic scatter within uncertainties. The measured intrinsic scatters do not change when we leave the slope and intercept (a and b) as free parameters or fix them to the values in Equations 4.2 and 4.3. Interestingly, this level of scatter is slightly smaller than the intrinsic scatter of ~ 0.1 dex found in the gas-phase MZR (e.g. Yabe et al., 2012; Guo et al., 2016) for galaxies with $M_* \gtrsim 10^{9.5} M_\odot$ but comparable to the intrinsic metallicity scatter of $\sim 0.05 - 0.08$ dex in the fundamental metallicity relation or in the gas-phase MZR when the star formation rate is taken into account (e.g. Mannucci et al., 2010; Yates et al., 2012; S. J. Lilly et al., 2013). This is expected because the stellar metallicity is less affected by the current star formation rate.

To further investigate the source of the intrinsic scatter, we plot the MZRs derived from both local and $z \sim 0.4$ galaxies, color-coded by their formation redshifts, in Figure 4.8a. Ideally, this should yield the MZR of the star-forming galaxies at each formation redshift, which should not, or at most weakly, depend on when the galaxies were observed, i.e.,the observed redshifts. Remarkably, the figure shows that, at each fixed mass, the galaxies that formed earlier (red data points) generally have lower metallicities than galaxies that formed later (blue data points) do. This is as we expect from the evolution of the gas-phase MZR, if stars approximately adopt the metallicity from their birth clouds.

At this point, we conjecture that the stellar MZR does not only depend on galaxy mass, but also on the redshift that galaxy formed (or when the majority of stars formed when the SSP is not assumed.) If this is true, the dependence on galaxy formation time should be able to explain the observed evolution with observed redshift. To see this effect better, we plot the deviation of the measured metallicities from the best linear fit of $z \sim 0$ galaxies (Equation 4.2, the upper dashed line) as a function of the age of the universe at the formation of their stellar populations in Figure 4.8c. By subtracting off the mass-metallicity function from the observed metallicities, the figure shows the effect of the age of universe at galaxy formation on stellar metallicity when the dependence on mass is removed. We can clearly see that galaxies that formed earlier (older galaxies) offset toward lower metallicities, while galaxies that formed later (younger galaxies) offset toward higher metallicities.

However, the correlation with formation redshift (older galaxies have lower metallicities) is in the same direction as the age–metallicity degeneracy. Though some previous works have pointed out the anti-correlation between age and metallicity at a given mass, it was complicated by or was thought to be the result of the age–metallicity degeneracy (e.g., Jørgensen, 1999; Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti, 2005). To test whether the scatter in the MZR as a function of formation redshift is real or caused by the age–metallicity degeneracy, we created a set of mock observed ages and metallicities. We assume that the metallicities are solely determined by stellar mass according to the best-fit linear function found in the observed MZR at $z \sim 0$. If the trend with age is caused by the age–metallicity degeneracy, then we should obtain the same level of scatter in the MZR and its correlation with formation redshift after noise is added.

For each observed galaxy, we construct its twin mock galaxy. We took the measured mass of each observed galaxy and calculated its "true" metallicity from the linear function found in Section 4.5 (Equations 4.2). We also took the measured age as the "true" age of the galaxy. The "true" observed spectrum was obtained from the FSPS according to its "true" age and metallicity, smoothed to the observed velocity dispersion convolved with SDSS/DEIMOS instrumental dispersion. Gaussian noise was added at each pixel with the same flux uncertainty array from the observed spectrum. We calculated a χ^2 grid for each noised spectrum with noiseless SPS spectra of every possible age and metallicity combination in the grid of 0.5 to 13 Gyr in age and -0.8 to 0.2 dex in metallicity. The grid spectra were smoothed to the same dispersion as the noised spectrum. The "observed" age and metallicity of

each mock spectrum was then selected according to the probability of each cell in the χ^2 grid. The resulting mock MZR is shown in Figure 4.8b.

We do not find the same level of separation of the mock MZR with galaxy formation redshift as in the observed MZR. The dots and diamonds of different colors in Figure 4.8b are visibly more mixed than those in Figure 4.8a. We plot the deviation of the "measured" metallicities from the best linear fit to the $z \sim 0$ stellar MZR as a function of the age of the universe at their formation (Figure 4.8d). The slope in Figure 4.8d (mock observation) is purely caused by the age–metallicity degeneracy. If there were no age–metallicity degeneracy, the deviation from the best-fit linear relation should scatter around Δ[Fe/H]=0 at all ages. However, the degeneracy causes the data points to move sligthly toward the lower left (more metal-poor and older) or upper right (more metal-rich and younger). Fitting a linear fit to the underlying probability distribution with a maximum likelihood estimation [3], we found a small positive slope of 0.005 ± 0.003 dex per Gyr in the relation in Figure 4.8d.

The slope in the relation between the observed deviation from the best-fit line in the MZR as a function of formation redshift is steeper than that of the mock observation. The slope of the deviation in Figure 4.8c is 0.055 ± 0.006 dex per Gyr, significantly larger than the slope in Figure 4.8d. When we fit linear functions to the $z \sim 0$ and the $z \sim 0.4$ sample individually, the best-fit slopes and intercepts are consistent within uncertainties. (The slopes are 0.058 ± 0.010 and 0.48 ± 0.07 dex per Gyr for the $z \sim 0$ and $z \sim 0.4$ sample, respectively.) This confirms that the evolution of the MZR with formation redshifts is real. Galaxies with SSP-equivalent formation redshift at $z \sim 2$ have [Fe/H] on average of 0.4 – 0.5 dex lower than the metallicities of galaxies that just formed in the past 2 Gyr.

The evolution of the MZR with formation redshift suggests that the mass–metallicity relation is not only determined by galaxy masses but also star formation histories. Moreover, this evolution of the MZR with formation redshift can consistently explain the evolution with observed redshift found in Section 4.5, which was 0.16 ± 0.03 dex from $z \sim 0.4$ to $z \sim 0$. The difference in the weighted mean formation time of

[3]Because the uncertainties of individual ages and metallcities are not Gaussian and highly correlate with each other (see Figure 4.5), we cannot use linear-fit estimators that assume Gaussian probability distribution. In this case, we use Markov chain Monte Carlo sampling to obtain the best linear fit that minimize the negative likelihood, $-\sum \log(\oint P_i dl(a,b))$, where P_i is the probability distribution of individual measurements of Δ[Fe/H] and age of the universe at galaxy formation. The integration is along the considered linear function with parameters a and b for the slope and intercept. The summation is over all data points. The best linear fit is the line that passes through the highest probability regions. Note that the underlying distribution in Figure 4.8c and 4.8d are $\sum P_i$, which represent uncertainties of the data.

the two populations is 2.7 ± 0.1 Gyr. If this difference is multiplied by the evolution with formation redshift 0.055 ± 0.006 dex per formation Gyr, we would expect a 0.15 ± 0.02 dex difference in [Fe/H] between the two populations, consistent with what we observed in the evolution with observed redshift. The gentler evolution of the MZR with observed redshift compared to the evolution with formation redshift is probably the result of a shared histories prior to quenching that smear out the evolution.

Lastly, we note that the evolution of the MZR with formation redshift seen here is inconsistent with the observations in dwarf galaxies. Kirby, Cohen, et al. (2013) measured metallicities of dwarf galaxies based on measurements of individual stars. The authors established that a single MZR applies to all Local Group galaxies with $10^{3.5} < M_*/M_\odot < 10^9$ regardless of their star formation histories. Dwarf irregular galaxies with gas present today have the same MZR as dwarf spheroidal satellites with no gas present today. If this scenario held true in the more massive population, we would expect a tight MZR regardless of formation redshift.

On the slopes of the MZR relations

Another interesting feature that emerges from the mock MZR is the curve that bends toward lower metallicities at the low-mass end, even though we constructed the mock MZR from a linear relation (black dashed line in Figure 4.8b). The curve is similar to the observed MZR. This suggests that the age–metallicity degeneracy causes the tendency to scatter toward lower metallicity in low-mass galaxies. An explanation can be found in the χ^2 contours in Figure 4.5. At [Fe/H]~ -0.2, approximately where the change of slope occurs, the contours of equal probabilities can be asymmetric, biasing toward lower metallicities than the true values. In addition, lower metallicities generally have larger uncertainties than the uncertainties at solar metallicities, causing larger scatter at the low-metallicity or the low-mass end. This finding could suggest that the stellar MZR might in fact be a single power law with a similarly tight dispersion, at least over the observed mass range of $M_* \approx 10^{9.7} M_\odot$ to $10^{11.5} M_\odot$. To confirm this, individual spectra of low-mass galaxies with high S/N are required to secure low uncertainties in [Fe/H], which is beyond our current work but might be achieved by using gravitationally lensed galaxies, larger telescopes, or longer exposure time.

The shallow slopes of ~ 0.16 dex per log mass found in both local and $z \sim 0.4$ MZRs may give new insight into the strength of feedback in galaxies of the observed mass

range. Strong star formation feedback generally results in a steep MZR (e.g. De Lucia et al., 2012; Lu, Wechsler, et al., 2014). Lu, Wechsler, et al. (2014) compared model predictions of both gas-phase and stellar MZRs from three independently developed semi-analytic models, namely the Croton model (Croton et al., 2006), the Somerville model (Somerville, Gilmore, et al., 2012; Porter et al., 2014), and the Lu model (Lu, Mo, et al., 2011). The authors found that the Croton model, which assumes a constant mass-loading factor, predicts the shallowest slope in the local stellar MZR, at ~ 0.17 dex per log mass over the mass range of $10^8 - 10^{11} M_\odot$. In contrast, the Lu model, in which the mass-loading factor is a strong function of halo circular velocity, predicts a very steep MZR, at ~ 0.6 dex per log mass over the same mass range. The slope of the MZR in our observations is consistent with the slope predicted from the Croton model.

Remarkably, our results agree with what Lu, Wechsler, et al. (2014) found based on the gas-phase MZRs at $z \lesssim 1$. Among the three models considered, the Croton model also describes the observations of gas-phase MZRs the best. This might suggest that, over the observed mass and redshift range, the amount of galaxy outflow is mainly a function of SFR and does not have a strong additional dependence on galaxy mass. This picture is closely related to the results from a simple closed box model and from the FIRE hydrodynamical simulations, where stellar metallicity is a strong function of gas fraction within a galaxy halo (Ma, Hopkins, Kasen, et al., 2016).

However, the slope of the stellar MZR found in this work is not consistent with the slope of the stellar-MZR found in dwarf galaxies. Dwarf galaxy satellites of the Milky Way exhibit a stellar MZR that is consistent with an unbroken power law (e.g. Grebel et al., 2003; Kirby, Cohen, et al., 2013). Based on measurements of metallicities of individual stars in local-group dwarf galaxies with masses ranging from $M_* \sim 10^3$ to $10^9 M_\odot$, Kirby, Cohen, et al. (2013) measured the slope of the MZR to be 0.30 ± 0.02 dex in [Fe/H] per log mass. The slope of the dwarf galaxy MZR is significantly larger than the slope found in this work, which is based on individual integrated spectra of both local and $z \sim 0.4$ galaxies with stellar masses $\gtrsim 10^{9.5} M_\odot$. If both numbers are correct, there must be a change of slope around the transition mass.

The slopes of our observed stellar MZRs do not change significantly when we limit the sample to those with lower masses, $M_* < 10^{10.5} M_\odot$. The slopes at this lower-mass end are 0.15 ± 0.04 and 0.33 ± 0.16 dex per log mass for $z \sim 0$ and $z \sim 0.4$

sample respectively. Although the slope at the low-mass end of the $z \sim 0.4$ sample may seem to suggest a change in slope, it is still consistent with the slope we found for the whole sample in Section 4.5. Furthermore, based on what we found from the mock data (Figure 4.8b), the age–metallicity degeneracy can cause the metallicity at the lower-mass end to bias low and create a seemingly steeper of slope.

Although Ma, Hopkins, Kasen, et al. (2016) fit a single power law to the MZRs of simulated galaxies over an entire 8 dex in mass, their FIRE hydrodynamical simulations indeed seem to show a change of slope around $M_* \sim 10^{8.5} M_\odot$ for the MZR of $z \sim 1.4 - 4$ simulated galaxies. Dwarf galaxies exhibit a steeper slope than higher-mass galaxies (as shown in Ma, Hopkins, Kasen, et al.'s Figure 4). Unfortunately, the sample sizes of the simulated galaxies at $z = 0$ and $z = 0.8$ are also not large enough to exhibit a clear change of slopes.

As discussed earlier, the slopes of the MZRs reflect the strength of mass loading factors. If the change of slope is real, the mass at the change of slope can suggest a mass below which feedback starts to have additional dependence on other parameters. In fact, Lu, Benson, et al. (2017) argued that at the low-mass regime, two different feedback mechanisms, i.e., ejective and preventive feedback, are needed to explain both the observed mass–metallicity relation and the stellar mass function.

4.6 Effect of Galaxy Environment

Observing galaxies in galaxy clusters have a benefit of being able to obtain multiple spectra in a few telescope pointings. However, we have so far ignored the fact that our samples at $z \sim 0.4$ are in a dense cluster environment and treated them as if they were general early-type galaxies. Here we discuss the impact of environment on our results.

Not all properties of galaxies have been shown to correlate with environment. Group and cluster environments show a higher fraction of passive galaxies than that of the field environment (e.g. Gerke et al., 2007; Muzzin, Wilson, et al., 2012; Koyama et al., 2013). At $z = 0.4$, based on the Hyper Suprime-Cam survey, the red fraction in cluster environments (number of members > 25) is about 40% higher in $M_* \sim 10^{9.5} M_\odot$ galaxies and about 20% higher in $M_* \sim 10^{10.75} M_\odot$ galaxies (Jian et al., 2018).

However, neither galaxy size nor the galaxy stellar mass function (GSMF) seems to depend on global environment predominantly. Morishita et al. (2017) found no significant differences in half-light radii between cluster or field systems in

the *Hubble Frontier Fields*. The shapes of the GSMFs of the general field and clusters are also mostly indistinguishable. The main difference is among the galaxies with $M_* \gtrsim 10^{11} M_\odot$, which are more enhanced in high-density environments (e.g. Calvi et al., 2013; Malavasi et al., 2017; Etherington et al., 2017). Therefore, the completeness of our sample relative to the cluster population should be more or less transferable to the completeness of the general population.

In terms of chemical composition, gas-phase metallicities in star-forming galaxies have been shown to have slight or no correlation with environment. More metal-rich galaxies, on average, reside in over-dense regions (e.g. Cooper, Tremonti, et al., 2008; Wu et al., 2017). Cooper, Tremonti, et al. (2008) used strong emission lines to measure gas-phase metallicities of SDSS star-forming galaxies. The authors found that the offset in metallicity relative to the median gas-phase MZR as a function of galaxy overdensity is significant. However, the metallicity offset between the least dense and densest environment considered in that study is less than 0.03 dex. For higher redshift galaxies, Kacprzak et al. (2015) studied gas-phase metallicities of star-forming galaxies in a $z \sim 2$ galaxy cluster. They found no distinguishable difference between the gas-phase MZR of field and cluster galaxies to within 0.02 dex.

The effects of environment on age and metallicity are also likely to be minimal for early-type galaxies. In particular, Fitzpatrick et al. (2015) found that SDSS quiescent early-type galaxies have slight variations in age with environment. Isolated galaxies have the youngest ages; brightest cluster galaxies are 0.02 dex older; and satellites are 0.04 dex older than the isolated galaxies. There is no significant variation in Fe enrichment. Furthermore, Harrison et al. (2011) measured ages and metallicities of early-type galaxies residing in four local galaxy clusters and their surroundings extending to 10 $R_{\rm vir}$. The ages and metallicities were measured via spectrophotometric indices. They found no dependence of age or metallicity on the locations of galaxies in the clusters, i.e., those in the clusters or in the clusters' outskirts. Harrison et al. concluded that galaxy mass plays a major role in determining stellar populations.

In conclusion, cluster environment can affect the chemical abundance in galaxies. However, the effect seems to be small and weaker in stellar metallicities than in gas-phase metallicities.

4.7 Summary

In this paper, we measured ages and metallicities of 62 individual quiescent galaxies in the $z \sim 0.4$ galaxy cluster Cl0024+17. The quiescent galaxies were selected based on the EW of the [O II] λ3727Å emission line and FUV−V color. The final sample spans the stellar mass range from $10^{9.7}$ to $10^{11.5} M_{\odot}$ with $\sim$ 50% completeness for $M_* \gtrsim 10^{9.7} M_{\odot}$. We employed a full spectrum fitting technique by adopting FSPS models (Conroy, Gunn, and M. White, 2009) and the assumption of single stellar populations (SSPs). We examined the accuracy of our fitting technique in several aspects, including varying the signal-to-noise ratios, testing the validity of the SSP assumption, and comparing with previous measurements in the literature. Our age and metallicity measurements have typical uncertainties of < 0.15 dex. We also measured ages and metallicities from a subsample of local SDSS quiescent galaxies from Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) and constructed the MZR of the local quiescent galaxies based on our measurements. We used this local MZR to compare with the MZR of $z \sim 0.4$ galaxies. We find the following:

1. We considered three functions (logarithmic, quadratic, and linear) to fit the MZRs at both redshifts. We found that the linear function fits the observed MZRs the best.

2. We detect an evolution of the stellar MZR with observed redshift at $> 5\sigma$. The evolution is 0.16 ± 0.03 dex from $z \sim 0.4$ to $z \sim 0$ or 0.037 ± 0.007 dex per Gyr. The observed evolution is greater but consistent within 1σ and 2σ uncertainties with predictions from the FIRE and the EAGLE hydrodynamical simulations. Our results may have emphasized the importance of recycling processes in hydrodynamical simulations.

3. The intrinsic scatter of the MZR is smaller than that of the gas-phase MZR but comparable to the scatter in the fundamental metallicity relation. The intrinsic scatter can be explained by an evolution of the MZR with galaxy age or formation redshift. The MZR of galaxies that formed earlier offsets toward lower metallicity in the same manner as the evolution of the gas-phase MZR. The offsets are significant and not caused by the age–metallicity degeneracy. The evolution of the MZR with formation time is 0.055 ± 0.006 dex per Gyr, which is stronger than and can explain the evolution with observed redshift.

4. Based on constructing a mock MZR from a linear relation, the age–metallicity degeneracy can cause the MZR at the low-metallicity end to offset to lower metallicity, creating a downward curve sometimes seen in measurements of the MZR.

5. The slope of the MZR is $\sim 0.16 \pm 0.03$ dex per log mass. The slopes are consistent with the predicted slope from a semi-analytic model (the Croton model, Croton et al., 2006) in Lu, Wechsler, et al. (2014), which employs a modest, galaxy mass-independent mass-loading factor. Our results suggest that galaxy feedback (in terms of mass-loading factor) might not have a strong additional dependence on galaxy mass over the observed mass and redshift range.

We will investigate the evolution of the MZR further in the next chapter, using galaxies at higher observed redshifts. We will also measure α enhancements, as indicators of the star formation timescales of those population.

The authors thank the referee for a constructive and helpful report. We also thank Anna Gallazzi and Jieun Choi for kindly providing data, catalogs, and feedback to the paper. Thank Xiangcheng Ma, Yu Lu, Shea Garrison-Kimmel and Robyn Sanderson for useful discussions. E.N.K. and N.L. acknowledge support from the National Science Foundation grant AST-1614081.

Table 4.1: Catalog of Measured Ages and [Fe/H]'s

No.	RA	DEC	log M_*	[Fe/H]	Age	S/N	No.	RA	DEC	log M_*	[Fe/H]	Age	S/N
1	00 25 51.07	+17 08 42.4	10.8	$+0.10^{+0.08}_{-0.10}$	$2.4^{+0.3}_{-0.2}$	16.0	32	00 26 31.73	+17 12 24.1	11.1	$-0.09^{+0.12}_{-0.07}$	$3.3^{+1.1}_{-0.8}$	15.4
2	00 25 54.52	+17 16 26.4	10.4	$-0.16^{+0.19}_{-0.17}$	$2.9^{+2.3}_{-1.0}$	9.0	33	00 26 31.81	+17 11 56.9	10.4	$-0.02^{+0.08}_{-0.08}$	$7.1^{+1.7}_{-1.0}$	8.0
3	00 25 57.73	+17 08 01.5	10.4	$-0.12^{+0.08}_{-0.09}$	$5.7^{+1.5}_{-2.0}$	8.4	34	00 26 32.50	+17 10 26.0	10.9	$-0.07^{+0.12}_{-0.20}$	$2.5^{+0.8}_{-0.3}$	10.6
4	00 26 04.44	+17 20 00.6	10.5	$-0.22^{+0.39}_{-0.17}$	$2.0^{+0.7}_{-0.8}$	8.6	35	00 26 32.71	+17 07 56.0	10.3	$-0.20^{+0.13}_{-0.27}$	$2.0^{+0.4}_{-0.4}$	10.0
5	00 26 05.80	+17 19 19.0	10.9	$-0.05^{+0.07}_{-0.10}$	$5.4^{+2.1}_{-2.6}$	11.9	36	00 26 33.54	+17 09 23.9	10.6	$-0.11^{+0.14}_{-0.60}$	$1.3^{+0.7}_{-0.1}$	19.3
6	00 26 05.84	+17 19 19.0	10.9	$-0.23^{+0.18}_{-0.20}$	$6.0^{+4.3}_{-2.2}$	11.5	37	00 26 33.60	+17 09 20.2	10.5	$-0.13^{+0.09}_{-0.07}$	$10.0^{+1.4}_{-2.2}$	8.1
7	00 26 06.95	+17 19 42.8	10.8	$-0.14^{+0.06}_{-0.05}$	$4.6^{+0.7}_{-1.0}$	25.7	38	00 26 33.66	+17 09 31.0	10.5	$-0.07^{+0.07}_{-0.07}$	$4.7^{+1.1}_{-1.7}$	12.3
8	00 26 08.85	+17 09 54.7	10.7	$-0.14^{+0.12}_{-0.14}$	$2.9^{+1.0}_{-0.4}$	18.3	39	00 26 33.81	+17 12 16.6	10.6	$-0.02^{+0.15}_{-0.07}$	$3.4^{+1.5}_{-0.6}$	8.4
9	00 26 09.66	+17 11 13.5	10.4	$-0.42^{+0.15}_{-0.15}$	$7.5^{+1.9}_{-3.2}$	8.1	40	00 26 34.35	+17 10 22.1	11.1	$-0.16^{+0.07}_{-0.07}$	$4.1^{+0.7}_{-1.0}$	27.4
10	00 26 13.92	+17 13 34.9	10.1	$-0.33^{+0.24}_{-0.17}$	$2.4^{+0.7}_{-0.4}$	8.1	41	00 26 34.59	+17 10 16.4	10.8	$-0.25^{+0.10}_{-0.15}$	$5.6^{+2.4}_{-1.0}$	17.9
11	00 26 15.16	+17 18 15.6	10.4	$-0.08^{+0.16}_{-0.31}$	$2.7^{+2.2}_{-0.5}$	10.8	42	00 26 34.98	+17 10 21.3	10.5	$+0.04^{+0.09}_{-0.06}$	$4.0^{+2.3}_{-1.7}$	16.4
12	00 26 18.45	+17 07 01.1	10.3	$-0.27^{+0.18}_{-0.12}$	$4.0^{+3.0}_{-1.0}$	9.0	43	00 26 35.70	+17 09 43.1	11.5	$+0.01^{+0.06}_{-0.03}$	$5.6^{+1.0}_{-0.6}$	27.8
13	00 26 21.49	+17 14 11.8	10.6	$-0.08^{+0.06}_{-0.08}$	$5.5^{+1.5}_{-2.0}$	13.3	44	00 26 36.77	+17 09 28.7	10.1	$-0.23^{+0.12}_{-0.15}$	$4.0^{+2.3}_{-1.1}$	10.5
14	00 26 22.90	+17 12 31.4	11.1	$+0.20^{+0.04}_{-0.09}$	$2.8^{+0.2}_{-0.1}$	23.7	45	00 26 37.27	+17 10 00.2	10.6	$+0.10^{+0.10}_{-0.12}$	$5.6^{+3.1}_{-1.7}$	15.3
15	00 26 22.91	+17 12 31.3	11.1	$-0.05^{+0.10}_{-0.08}$	$4.3^{+1.6}_{-2.0}$	9.9	46	00 26 37.52	+17 09 08.7	10.6	$-0.18^{+0.07}_{-0.07}$	$6.9^{+1.0}_{-0.9}$	16.8
16	00 26 24.82	+17 12 21.5	10.9	$-0.14^{+0.11}_{-0.30}$	$5.2^{+5.2}_{-2.0}$	12.3	47	00 26 37.90	+17 09 22.0	11.2	$-0.01^{+0.05}_{-0.06}$	$4.8^{+1.5}_{-0.8}$	26.3
17	00 26 26.11	+17 11 57.8	10.3	$-0.02^{+0.15}_{-0.15}$	$5.4^{+3.2}_{-3.6}$	8.3	48	00 26 37.91	+17 09 37.8	10.7	$-0.07^{+0.09}_{-0.05}$	$3.8^{+1.2}_{-0.8}$	14.5
18	00 26 27.12	+17 12 25.9	10.9	$-0.16^{+0.12}_{-0.15}$	$9.0^{+2.7}_{-3.1}$	12.1	49	00 26 38.41	+17 09 58.7	10.3	$-0.14^{+0.13}_{-0.09}$	$3.2^{+1.1}_{-0.6}$	10.7
19	00 26 27.98	+17 11 37.7	10.5	$-0.30^{+0.12}_{-0.22}$	$8.5^{+3.1}_{-2.5}$	8.2	50	00 26 38.65	+17 09 14.5	9.7	$-0.15^{+0.17}_{-0.28}$	$3.9^{+4.4}_{-2.9}$	10.9
20	00 26 29.11	+17 10 24.7	10.7	$-0.07^{+0.07}_{-0.04}$	$3.7^{+1.0}_{-0.7}$	18.0	51	00 26 38.80	+17 09 59.5	10.7	$-0.11^{+0.21}_{-0.18}$	$2.7^{+2.3}_{-1.0}$	24.5
21	00 26 29.50	+17 10 32.6	10.5	$-0.15^{+0.12}_{-0.14}$	$3.1^{+1.2}_{-0.5}$	18.3	52	00 26 40.13	+17 08 21.5	9.9	$-0.60^{+0.08}_{-0.09}$	$4.6^{+0.7}_{-0.9}$	8.7
22	00 26 29.92	+17 10 06.8	10.3	$-0.21^{+0.04}_{-0.21}$	$1.3^{+0.0}_{-0.1}$	16.4	53	00 26 41.16	+17 10 01.8	11.1	$+0.01^{+0.19}_{-0.25}$	$3.0^{+2.7}_{-1.4}$	16.0
23	00 26 30.08	+17 07 49.1	9.8	$-0.21^{+0.29}_{-0.26}$	$2.1^{+0.8}_{-0.5}$	7.9	54	00 26 43.23	+17 08 41.0	10.1	$-0.25^{+0.36}_{-0.18}$	$4.0^{+2.2}_{-5.1}$	11.5
24	00 26 30.76	+17 12 26.3	10.3	$-0.15^{+0.13}_{-0.22}$	$1.6^{+0.6}_{-0.3}$	13.2	55	00 26 43.70	+17 07 12.8	10.0	$-0.06^{+0.15}_{-0.17}$	$3.0^{+1.5}_{-0.4}$	8.2
25	00 26 31.00	+17 17 09.0	10.0	$+0.10^{+0.06}_{-0.10}$	$2.3^{+0.3}_{-0.2}$	17.7	56	00 26 44.68	+17 08 33.8	9.9	$-0.79^{+0.35}_{-0.13}$	$6.8^{+1.6}_{-0.9}$	8.3
26	00 26 31.04	+17 11 09.3	10.4	$-0.09^{+0.08}_{-0.22}$	$4.0^{+2.7}_{-1.7}$	12.5	57	00 26 48.22	+17 10 46.3	10.2	$-0.27^{+0.06}_{-0.16}$	$1.7^{+0.2}_{-0.2}$	8.3
27	00 26 31.22	+17 12 08.4	10.7	$+0.06^{+0.13}_{-0.08}$	$3.1^{+1.7}_{-1.1}$	21.3	58	00 26 48.30	+17 12 35.4	10.6	$-0.28^{+0.05}_{-0.18}$	$1.3^{+0.0}_{-0.1}$	16.2
28	00 26 31.40	+17 17 00.3	11.1	$-0.06^{+0.14}_{-0.17}$	$5.0^{+2.5}_{-3.3}$	13.5	59	00 26 51.84	+17 08 39.9	10.5	$-0.21^{+0.10}_{-0.20}$	$6.5^{+2.8}_{-1.6}$	15.7
29	00 26 31.41	+17 10 55.6	11.3	$-0.15^{+0.04}_{-0.06}$	$6.3^{+0.7}_{-0.8}$	20.9	60	00 26 54.38	+17 08 27.0	11.0	$+0.10^{+0.11}_{-0.09}$	$3.5^{+2.8}_{-1.4}$	24.8
30	00 26 31.41	+17 10 27.1	10.0	$-0.27^{+0.13}_{-0.16}$	$5.0^{+1.4}_{-3.5}$	8.0	61	00 27 07.76	+17 10 49.5	11.0	$+0.10^{+0.10}_{-0.13}$	$3.0^{+1.0}_{-0.3}$	24.8
31	00 26 31.56	+17 17 12.8	10.5	$-0.12^{+0.04}_{-0.06}$	$5.9^{+0.7}_{-1.3}$	22.5	62	00 27 25.16	+17 07 22.4	10.1	$-0.51^{+0.24}_{-0.18}$	$6.1^{+1.7}_{-3.4}$	12.0

The units of M_*, age, and S/N are $M_\odot$, Gyr, and Å^{-1}, respectively.

APPENDIX

4.A Comparisons with existing measurements

We compare our measurements with literature measurements of the same galaxies (Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti, 2005; Choi, Conroy, et al., 2014). We selected a subsample from 44,254 SDSS spectra from Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) that is comparable to our $z = 0.4$ sample based on emission line EWs and U−B colors. Since the SDSS spectra cover up to at least 8000 Å in the rest frame but do not necessarily include the [O II]λ3727 Å emission lines, we instead used the criterion of rest-frame Hα EW < 1 Å to define quiescent galaxies. The limit of Hα EW = 1 Å was chosen so that the SFR is comparable to the SFR when the [O II] EW is equal to 5 Å. These limits were based on the SFR calibrations from Kewley, Geller, et al. (2004) assuming color $B - V = 2$, a typical color limit for quiescent galaxies (e.g. Schawinski et al., 2014) and no dust extinction. The color cut of U−B > 1 is to make sure that the contamination from star-forming galaxies is minimized (Mendez et al., 2011) in a similar manner to the color cut in our $z \sim 0.4$ sample. We selected all but at most 8 random quiescent early-type galaxies from each bin of 0.1 dex in logarithmic mass spanning the stellar mass range from 10^9 to $10^{11.5} M_\odot$. This sums to a subsample of 155 quiescent galaxies. To be consistent with the observed $z \sim 0.4$ spectra, we limited the wavelength range of the SDSS spectra to 3700–5500 Å. We repeated the age and metallicity measurements in the same manner as in Section 4.3.

Our measured metallicities agree reasonably well with the values measured by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005). The results

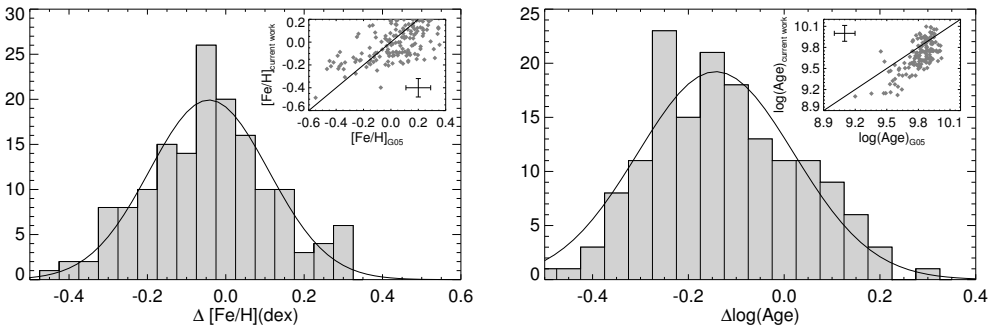

Figure 4.9: Comparison between the age and metallicities measurement in this work and Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) for the same subsample of SDSS quiescent galaxies.

are shown in Figure 4.9. The differences in the [Fe/H] measurements follow a Gaussian distribution with a width of ~ 0.15 dex, peaking at -0.04 dex. Though the width of the metallicity differences is comparable to the typical measurement uncertainty of 0.12 dex quoted by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005), there are some systematic differences in the measured [Fe/H]. The corner plot in Figure 4.9 shows that the metallicities we measured are slightly lower than those measured by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) at high metallicities but the offsets reverse at lower metallicities. We argue that the main reasons for the discrepancy are the differences in the stellar libraries used in generating model spectra, which will be discussed below together with the uncertainties in age measurements.

The differences in age measurements show larger discrepancies than those of metallicities. The distribution of differences in age peaks at ~ -0.15 to -0.25 dex with a Gaussian width of 0.16 dex. The galaxies were generally younger than reported by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) by about 0.2 dex. The discrepancy is likely SPS-model dependent. Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) computed 5 spectrophotometric indices from the BC03 stellar population synthesis, which is based on the STELIB spectral library (J.-F. Le Borgne et al., 2003). The FSPS models used in the current work are based on the MILES spectral library (Sánchez-Blázquez et al., 2006). Although both spectral libraries are empirical, the STELIB library contains fewer stars - 249 stellar spectra as compared to 945 spectra in the MILES library. Very few stars in the STELIB library are at non-solar metallicities (Conroy and Gunn, 2010).

Koleva et al. (2008) compared three spectral synthesis models of single stellar populations. In particular, the authors inverted the parameters from the SSP spectra produced by BC03 using a grid of models made with the Vazdekis/MILES (Vazdekis, Sánchez-Blázquez, et al., 2010) and the Pegase-HR spectral synthesis model (D. Le Borgne et al., 2004). The two models are based on the MILES library and the ELODIE library (Prugniel et al., 2007), respectively. They found that at approximately solar metallicity, the ages retrieved by both Vazdekis/MILES and Pegase-HR are ~ 0.2 dex younger than the input ages in BC03 when the input age is greater than ~ 5 Gyr or $\log(\mathrm{Age/yr}) \gtrsim 9.7$ (see Koleva et al., 2008, Figure 2a). The trend reverses at $\sim 0.2 - 0.3$ dex above the solar metallicity or 0.5 dex below the solar metallicity. Our finding is somewhat consistent with this result, in which we measured the ages to be $\sim 0.15 - 0.25$ dex younger than measured by Gallazzi,

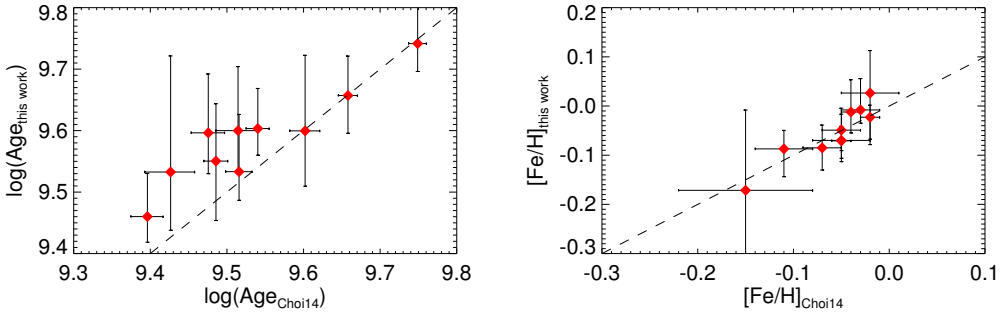

Figure 4.10: Comparison between the age and metallicities measurement in this work and Choi, Conroy, et al. (2014) for the same set of stacked spectra.

Charlot, Brinchmann, S. D. M. White, and Tremonti (2005), as shown in Figure 4.9.

Lastly, we found good agreement between our measurements and those by Choi, Conroy, et al. (2014) in both ages and metallicities. We obtained a set of stacked spectra compiled by Choi, Conroy, et al. (2014)[4]. These spectra were stacked from individual spectra of galaxies observed by the AGN and Galaxy Evolution Survey (AGES; Kochanek et al., 2012), in bins of redshifts from $z = 0.3$ to 0.7 and masses from $M_* = 10^{10.2}$ to $10^{11.3} M_\odot$. The spectral resolution is 6 Å, roughly double the resolution of our DEIMOS data. The wavelength coverage of the AGES spectra is 4000 to 5500 Å, which does not cover the full age-sensitive Balmer break region.

We found that our measurements of [Fe/H] are consistent with the values measured by Choi, Conroy, et al. (2014) (see Figure 4.10). For ages, all the values are consistent within 0.1 dex. However, there are higher discrepancies in the populations younger than 3.5 Gyr old, in the sense that we measured the ages to be slightly older than reported by Choi, Conroy, et al. (2014). The trend of this discrepancy is opposite to the discrepancy found when we compared our age measurements with Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005).

Discrepancies between our work and Choi, Conroy, et al. (2014) likely arise from differing wavelength coverage and the spectral models used. Because the AGES spectra only cover from 4000 to 5000 Å, higher order Balmer lines, which contain age information, are not present in the spectra. Moreover, Choi, Conroy, et al. (2014) used the SPS model from Conroy and van Dokkum (2012, p. CVD12). The main difference between the FSPS and the CVD12 model is that the latter allows abundances to be at non-solar ratios. Therefore, Choi, Conroy, et al. (2014) fit the

[4]Kindly obtained via private communication.

spectra for abundances of individual elements including [Mg/Fe], [O/Fe], [C/Fe], [N/Fe], etc. Though it would be beneficial to measure individual elements in our observed spectra, using the solar metallicity models without adjusting individual elements can still provide a reasonable fit to the spectrum (Conroy and van Dokkum, 2012). Besides, the CVD12 is not yet readily applicable to our data at lower masses because the model is limited to fairly a small range around solar metallicity ([Fe/H]$\in (-0.4, 0.4)$) and age greater than 3 Gyr. In fact, the lower age limit at 3 Gyr in the CVD12 might be responsible for the small age discrepancies found in Figure 4.10.

Chapter 5

EVOLUTION OF THE STELLAR MASS–METALLICITY RELATION. II. CONSTRAINTS ON GALACTIC OUTFLOWS FROM THE MAGNESIUM ABUNDANCES OF QUIESCENT GALAXIES

Leethochawalit, N., Kirby, E. N., Ellis, R. S., Moran, S. M., & Treu, T., Submitted to ApJ

Abstract

We present the stellar mass–[Fe/H] and mass–[Mg/H] relation of quiescent galaxies in two galaxy clusters at $z \sim 0.39$ and $z \sim 0.54$. We derive the age, [Fe/H], and [Mg/Fe] for each individual galaxy using a full-spectrum fitting technique. By comparing with the relations for $z \sim 0$ SDSS galaxies, we confirm our previous finding that the mass–[Fe/H] relation evolves with redshift. The mass–[Fe/H] relation at higher redshift has lower normalization and possibly steeper slope. However, based on our sample, the mass–[Mg/H] relation does not evolve over the observed redshift range. We use a simple analytic chemical evolution model to constrain average outflow that these galaxies experience over their lifetime, via the calculation of mass-loading factor. We find that the average mass-loading factor η is a power-law function of galaxy stellar mass, $\eta \propto M_*^{-0.21\pm0.09}$. The measured mass-loading factors are consistent with the results of other observational methods for outflow measurements and with the predictions where outflow is caused by star formation feedback in turbulent disks.

5.1 Introduction

Over the past five decades, we have made significant progress in measuring gas metallicities in star-forming galaxies. The relation between galaxy luminosity (stellar mass) and interstellar oxygen abundances in extragalactic H II regions was established 50 years ago (e.g. McClure et al., 1968; Lequeux et al., 1979; Garnett and Shields, 1987). Tremonti et al. (2004) used the large statistical sample size of local star-forming galaxies from the Sloan Digital Sky survey (SDSS) and confirmed that the correlation between metallicity and mass is more fundamental than that between metallicity and luminosity. The mass–metallcity relation (MZR) is such that more

massive galaxies have higher gas metallicities. In later work, the gas-phase MZR was also found to be present at high redshift and also to evolve with redshift (e.g. Maiolino, Nagao, et al., 2008; Zahid, Geller, Kewley, et al., 2013).

Although the gas-phase MZR has been known for almost four decades, its physical drivers are still debated. Early works mostly suggested that galactic winds are the primary agent that drives the relation (e.g., Mathews et al., 1971; Larson, 1974; Garnett, 2002; Tremonti et al., 2004). Lower mass galaxies have shallower potential wells and therefore can retain less of the metals they produced (Dekel et al., 1986). However, later works argued that the metallicity is regulated by a more complex mechanism that includes the interplay between inflow, outflow, and enrichment rate (e.g. Finlator et al., 2008; Davé, Finlator, et al., 2011). Spitoni et al. (2010) attempted to use chemical evolution models to explain the observed gas-phase MZR in the SDSS galaxies. They found that a range of models that include either outflow only, both inflow and outflow, or inflow with variable outflow can all explain the observed data equally well. In essence, there is a degeneracy between inflow and outflow rate that cannot be differentiated using only the data from the gas-phase MZR.

However, metals do not reside only in gas. Especially at lower redshift $z < 0.5$, the cold gas fraction is $< 20\%$ in most star-forming galaxies and less than a few percent in quiescent galaxies (e.g. Gobat et al., 2018). The majority of disk and metal mass is in stars (e.g., Werk et al., 2014). When Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) measured the stellar metallicities of local SDSS galaxies using spectral indices, they found that stellar metallicity also exhibits a tight correlation with stellar mass in galaxies with $M_* > 10^9\ M_\odot$. Kirby, Cohen, et al. (2013) measured stellar metallicities of individual stars in the Local Group and found that the correlation extends down to dwarf galaxies with stellar mass as low as $10^3\ M_\odot$.

In principle, measuring stellar metallicity in addition to gas-phase metallicity can break the degeneracy between inflow and outflow (e.g., Lu, Blanc, et al., 2015), resulting in better constraints on chemical evolution models. However, recent works that have attempted to incorporate both stellar and gas-phase MZRs found that it is difficult to reconcile the two MZRs. Lian et al. (2018) found that the stellar metallicities of local galaxies are generally lower than expected based on their gas-phase metallicities. The discrepancies are larger for lower mass galaxies. The only models that can reconcile both MZRs have to invoke either a steep initial-mass

function (IMF) slope (almost twice the slope of the Salpeter 1955 IMF) or a strong outflow (ejection of all metals produced) at early times.

Determining stellar metallicities in star-forming galaxies is challenging and subject to potentially large biases. First, emission lines from the interstellar medium have to be either subtracted or modeled together with stellar absorption lines. Second, even if the emission lines are modeled perfectly, the measurements of age and metallicity are still subject to bias, especially in the spectra of young stellar populations. The bias generally makes the population appear older and more metal poor than the true values (e.g., Leethochawalit, Kirby, et al., 2018) by as large as ~ 0.5 dex if the priors are not treated carefully (Ge et al., 2018; Cid Fernandes, 2018).

Quiescent galaxies provide an alternative and a more convenient way to constrain galaxy chemical evolution models. The populations are older. The contamination from emission lines is less concerning. The spectral analysis is overall less prone to systematic biases. Additionally, the (near) absence of gas in quiescent galaxies simplifies chemical evolution models by eliminating the need to consider metals in the gas.

In this work, we continue the work of Chapter 4 in quantifying the stellar MZR of quiescent galaxies as a function of redshift. We expand the sample size, from galaxies in a galaxy cluster at $z = 0.39$ to include an additional sample in a galaxy cluster at $z = 0.54$. We measure their magnesium [Mg/H] abundances in addition to iron [Fe/H] abundances. The highlight of this work is that we propose an archaeological method to constrain galactic outflow, in terms of average mass-loading factors that quiescent galaxies experienced, via the measurements of α element abundances. We build upon the technique introduced in Chapter 4 to trace quiescent galaxies back to their epochs of formation using their ages. We conclude that while the [Fe/H] abundance at a fixed galaxy mass appears to evolve with redshift, the [Mg/H] MZR changes neither with the redshift of formation nor with the redshift of observation.

5.2 Data

In Chapter 4, we presented a stellar MZR based on quiescent galaxies from the galaxy cluster Cl0024+17 at $z = 0.39$. In this paper, we expand the sample by including additional quiescent galaxies from the galaxy cluster MS0451 at $z = 0.54$.

The two galaxy clusters comprised the survey of Moran, R. S. Ellis, Treu, Smith, et al. (2007). The survey provides UV to near-infrared imaging and ground-based

optical spectroscopy of member galaxies up to $\sim$ 10 Mpc in diameter centered on both clusters. The available photometric bands are NUV and FUV from the Galaxy Evolution Explorer (GALEX) satellite, *BVRI* bands from the 3.6 m Canada-France-Hawaii telescope (Cl0024) and the Subaru 8 m telescope (MS0451), F814W ($\sim I$ band) from Hubble Space Telescope, and J and K_s bands from the WIRC camera on the Palomar/Hale 200" telescope. Using the same method as in Chapter 4, we use this photometry to estimate the stellar masses of the sample with the SDSS `KCORRECT` software version v4_3 (Blanton et al., 2007).

The spectroscopic data for both Cl0024 and MS0451 were obtained with the DEIMOS spectrograph (Faber, Phillips, et al., 2003) on the Keck II Telescope. Most data are part of the original survey. We obtained additional DEIMOS spectroscopy to enhance the S/N of a subset of the galaxy spectra in MS0451. In both clusters, the spectra were obtained with 1" wide slits with a spectral resolution of at least $R = 2000$, spanning rest-frame wavelengths from $\sim$ 3500 to 6000 Å. The detailed spectroscopic observations of Cl0024 are described by Moran, R. S. Ellis, Treu, Smail, et al. (2005) and Chapter 4.

We now summarize the previous and new DEIMOS spectroscopy in MS0451. In 2003, Moran, R. S. Ellis, Treu, Smail, et al. (2005) observed 11 slitmasks for 1 hr each with the 600 line mm^{-1} grating centered at 7500 Å. Based on these initial data, the study identified cluster members and performed a deeper follow-up in 2004–2005 with 10 additional slitmasks with the same grating centered at 6800 Å. The integration time was 2.5–4 hrs per slitmask. In total, the original survey identified 319 member galaxies in MS0451. Moran, R. S. Ellis, Treu, Smith, et al. (2007) gives further details of these observations. In December 2016 and October 2017, we additionally observed one more slitmask with the same grating centered at 7200 Å for a final follow-up. In the mask design, we prioritized previously observed quiescent galaxies whose spectra combined over all observing runs could achieve S/N of at least 8 Å^{-1} within 5 additional hours of integration time. The actual integration time was 4.5 hrs.

We selected the final sample in a manner similar to Chapter 4 with the following criteria. First, the galaxies are quiescent, which we defined as having equivalent widths (EWs) of [O II] λ3727 less than 5 Å and either having rest-frame FUV $- V$ colors larger than 3 or no detection in rest-frame FUV. Second, the signal to noise (S/N) is greater than 10 Å^{-1} in rest-frame. This resulted in the final sample of 59 galaxies in Cl0024 and 92 galaxies in MS0451. The lowest masses are $M_* = 10^{9.7}$

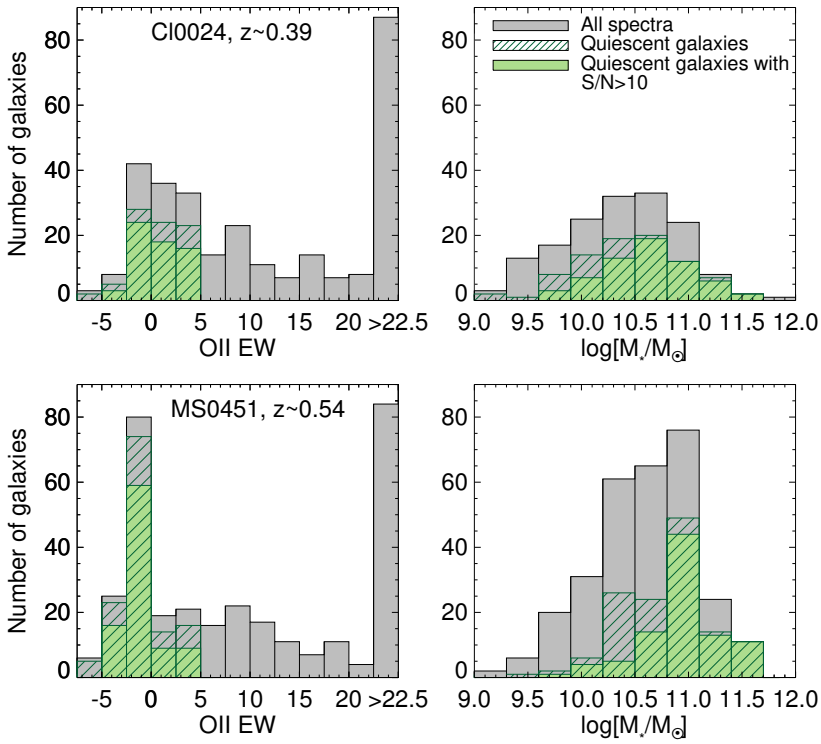

Figure 5.1: Histograms of the parent and selected sample in this paper. We select quiescent galaxies (defined in Section 5.2) with S/N greater than 10 Å^{-1} in the rest frame to be in our sample (solid green).

and $10^{9.6}$ $M_{\odot}$, respectively (see Figure 5.1).

All DEIMOS spectra were reduced using the spec2d DEIMOS data reduction pipeline (S. F. Newman, Genzel, Förster Schreiber, et al., 2013) adapted by Kirby, Simon, et al. (2015). Each spectrum was flat-fielded, wavelength-calibrated, sky-subtracted, and telluric-corrected.

Finally, we include a subsample of $z \sim 0$ SDSS quiescent galaxies in our sample to compare with our observed galaxies. We use the same 155 randomly selected SDSS quiescent galaxies from Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti's (2005) sample in the mass range 10^9 to $10^{11.5} M_{\odot}$ as in Chapter 4. To recapitulate, we put limits on broadband color and on the maximum Hα EW so that the sample selection criteria are close to what we used for the higher redshift sample. We select roughly equal numbers of SDSS galaxies in each mass bin.

5.3 Model Fitting

We use full spectral fitting to derive ages and stellar metallicities. As in Leethochawalit, Kirby, et al. (2018), we adopt the single stellar population (SSP) models from

the Flexible Stellar Population Synthesis (FSPS, Conroy, Gunn, and M. White, 2009) version 3.0. We generate the SSP spectra with the Kroupa (2001) IMF, Padova isochrones (Marigo and Girardi, 2007), and the MILES spectral library (Sánchez-Blázquez et al., 2006). The models have metallicity and age ranges of $-1.98 < \log Z < 0.2$ and 0.3 Myr to 14 Gyr, respectively. To calculate the effect on the integrated light spectra by specific elements, we use the theoretical response functions from Conroy, Villaume, et al. (2018), which depend on metellicity and age. They were computed from the Kurucz suite of theoretical model atmospheres and spectra (Kurucz, 1993).

We are interested in 3 parameters: age, [Fe/H], and [Mg/Fe]. We choose to measure [Mg/Fe] for two reasons. First, magnesium is one of the alpha elements that is mainly produced by Type II supernovae. Its recycling time can thus be approximated as instantaneous. Second, magnesium absorption features are distinct. Specifically, the Mg b absorption lines at 5170 Å minimally overlap with absorption features of other elements. This makes the measurement more reliable than other instantaneously recycled elements, especially in low-S/N spectra.

We assume that each galaxy is a single stellar population, i.e., all stars in the galaxy were born at the same time with the same metallicity. Thus, our obtained ages and metallicities are SSP-equivalent values. This means that the measured age is expected to be younger than both light- and mass-weighted ages derived in models with extended star-formation histories (Choi, Conroy, et al., 2014). The metallicities are less affected (Mentz et al., 2016).

In order to measure age, [Fe/H], and [Mg/Fe], we actually need to fit for 6 parameters, all of which influence the 3 parameters of interest: age, [Z/H], velocity dispersion, redshift, [Mg/Fe] and [N/Fe]. Although we are mainly interested in [Mg/Fe], we experimented with a few options for the combinations of additional metal enhancements in Appendix 5.A. We found that by fitting for two additional abundance ratios—[Mg/Fe] and [N/Fe]—we obtain the values of interest that are most consistent with literature and with the results from fitting with a more elaborate set of metals: [Mg/Fe], [O/Fe], [C/Fe], [N/Fe], [Na/Fe], [Si/Fe], [Ca/Fe], and [Ti/Fe].

We interpret the measured [Z/H] as [Fe/H] for the following reason. In our case, the measured [Z/H] is the metallicity of the base SSP model (without Mg and N enhancement) and therefore not the actual total metallicity of the galaxy. Each base SSP model is the integrated stellar spectrum of stars that lie on the isochrone of a

given [Z/H] and age. However, the Padova isochrones assume solar-scaled abundances of individual metal elements ([Fe/H] = [Z/H]). In addition, the metallicities in the MILES stellar library were measured in terms of [Fe/H]. Thus, our measured [Z/H] is best interpreted as [Fe/H]. Besides, based on results in Appendix 5.A, our measured [Fe/H] agrees well with [Fe/H] measured from more complex models.

The fitting method is the same as that in Chapter 4. In summary, we first mask out the most prominent telluric band in the 7591–7703 Å wavelength range. In addition, if the spectrum has a positive EW of the [O II]λ3727 emission line, we then also mask out the wavelength range of the [O II]λ3727, [O III]λ5007, 4959 and Hβ line. We iteratively fit each spectrum with the Levenberg–Marquardt χ^2-minimization method.

For each spectrum, we fit for at least 100 iterations with the IDL code MPFIT (Markwardt, 2012) until the fitting parameters converge. In the first iteration, we fit the continuum-normalized observed spectrum with continuum-normalized model spectra. In the subsequent iterations, we applied the continuum from the best-fit model of the previous iteration to the observed continuum-normalized spectrum. We then fit the resulting spectrum with the model spectra, including the continuum shape. Lastly, to avoid the convergence of parameters on local minima, we fit each spectrum at least 5 times with different initial parameters, and we adopt the results from the models with the least χ^2.

Finally, we convert the measured enhancement of the [Mg/Fe] from the base FSPS spectra to the actual [Mg/Fe]. We followed the method in Conroy, Villaume, et al. (2018) by adding the abundance pattern of the MILES library stars at a given [Fe/H] to the measured Mg enhancement to obtain the final [Mg/Fe]. Because the [Mg/Fe] pattern of the stellar library in the measured [Fe/H] range only varies within ~ 0.05 dex, this step does not significantly affect the reported value of [Mg/Fe].

We calculate the uncertainties of the measured ages and metal abundances from mock SSP spectra. We first generate a mock spectrum with the measured [Fe/H], [Mg/Fe], and age. We then smooth the spectrum to a fixed velocity of 250 km/s, convolve it with the observed spectral resolution, and add Gaussian noise to reach the same S/N of the observed spectrum. We lastly compare the noised spectrum to a 3-dimensional grid of noise-less SSPs (in age, [Fe/H] and [Mg/Fe]) and calculate the uncertainties based on the χ^2 grid.

Examples of best-fit models of the spectra at the lower and upper end of the S/N

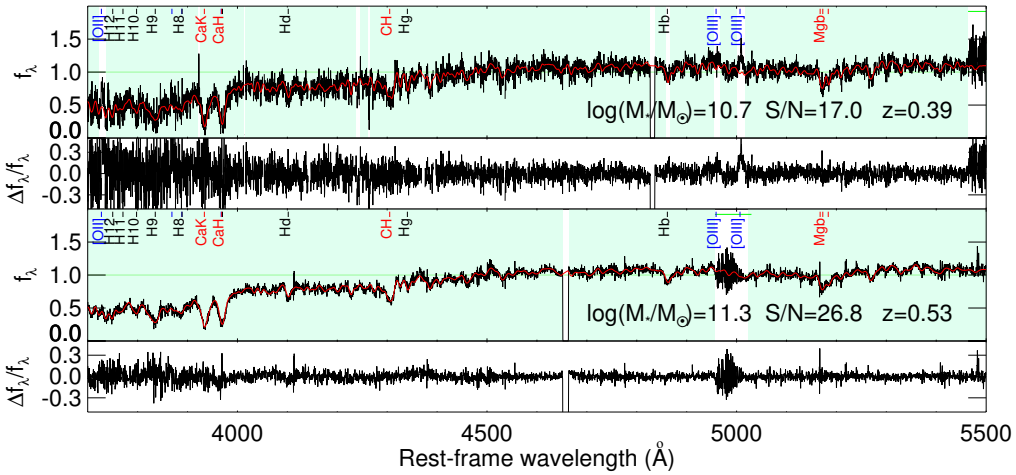

Figure 5.2: Examples of two observed spectra (black), best fit models (red), and the model residual (bottom panel of each spectrum). The top panels are an example of the spectra at the lower end of the S/N cut from a member of Cl0024 cluster. The bottom panels show a high S/N spectrum from the MS0451 cluster. The best-fit parameters are Age= $3.3^{4.7}_{2.6}$ and $4.3^{6.0}_{3.4}$ Gyr; [Fe/H]= $-0.0^{0.11}_{-0.7}$ and $-0.05^{0.03}_{-0.11}$ dex; [Mg/Fe]= $0.39^{0.52}_{0.28}$ and $0.39^{0.54}_{0.35}$ dex, respectively.

range are shown in Figure 5.2. The list of all measured parameters are in Table 5.2 and 5.3.

5.4 Results

The MZR measured from [Fe/H] evolves with redshift

We show the relation between [Fe/H] and galaxy stellar mass in Figure 5.3. Visually, the galaxies at higher redshift (Cl0024 and MS0451) appear to have lower [Fe/H] abundances than those of local SDSS galaxies, especially in the lower mass range. We use the analysis of covariance to quantify the dependence of MZR on redshift.

First, we assume that both normalization and slope of the mass–[Fe/H] relation depend on redshift. We fit data from every redshift at once with a single linear multiple regression, where [Fe/H] depends on one quantitative variable (mass) and categorical variables (different redshift samples). We choose to fit with a linear function because we found in Chapter 4 that both a two-degree polynomial and a logarithmic function introduce degrees of freedom that are not justified by the data.

The first model we consider (model 1), allows interactions[1] between mass and

[1] "Interaction" is a term in statistics that describes a situation where the effect of one independent variable depends on the magnitude of the other independent variable. In this case, we allow the evolution of metallicity with redshift to be mass-dependent.

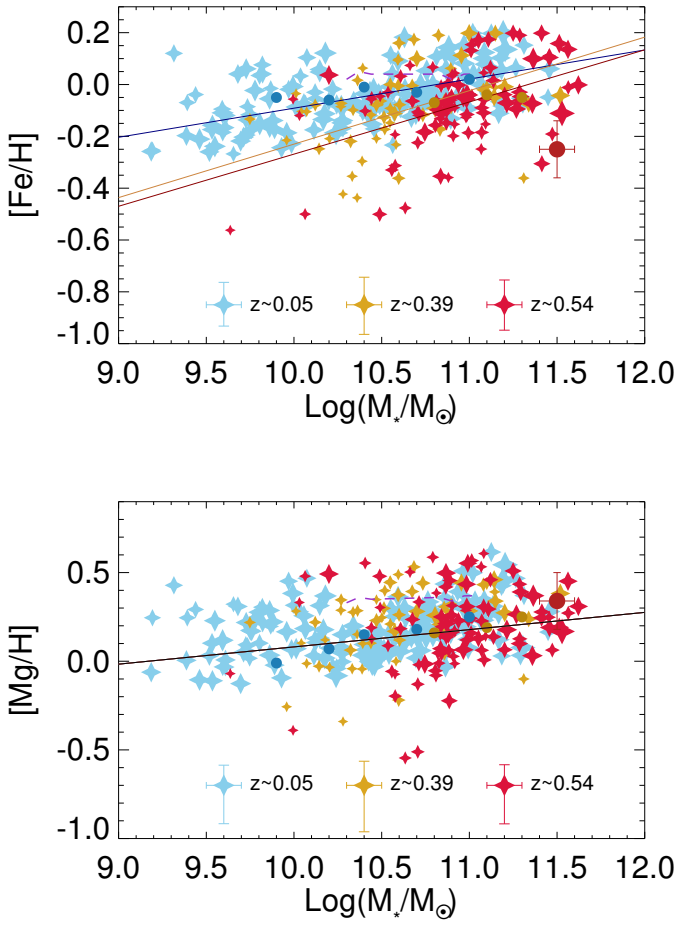

Figure 5.3: Stellar mass–metallicity relations measured from [Fe/H] (top) and [Mg/H] (bottom). Measurements in this paper are in diamonds. The labels indicate different redshift samples and their median uncertainties. Best-fit linear models from Table 5.1 are plotted in solid lines (model 1 for [Fe/H] and model 3 for [Mg/H]). Measurements based on stacked spectra of similar redshifts by Choi, Conroy, et al. (2014) are in solid circles (dark blue and dark yellow for z = 0.1–0.2 and z = 0.4–0.55 redshift bins, respectively). Magenta dashed lines are predictions based on the $z \sim 0$ quiescent population in the IllustrisTNG simulation (Naiman et al., 2018). We found a significant evolution with redshift in the [Fe/H]–mass relation, but not in the [Mg/H]–mass relation.

Table 5.1: Model parameters for MZRs

			[Fe/H] vs mass			[Mg/H] vs mass		
model	parameter/description		value	t	sig.	value	t	sig.
1	β_0	constant	-0.09 ± 0.01	7.9	<0.0005	0.09 ± 0.02	4.5	< 0.0005
	β_M	mass	0.11 ± 0.02	7.2	<0.0005	0.10 ± 0.03	3.8	< 0.0005
	β_{cl}	cl0024	-0.14 ± 0.05	2.9	0.002	-0.06 ± 0.07	0.8	0.2
	β_{ms}	ms0451	-0.18 ± 0.05	3.9	<0.0005	-0.01 ± 0.07	0.2	0.4
	$\beta_{cl\cdot M}$	cl0024*mass	0.09 ± 0.06	1.6	0.05	0.12 ± 0.09	1.4	0.08
	$\beta_{ms\cdot M}$	ms0451*mass	0.09 ± 0.04	2.0	0.02	-0.03 ± 0.07	0.4	0.4
2	β_0	constant	-0.10 ± 0.01	9.3	< 0.0005	0.09 ± 0.02	4.2	< 0.0005
	β_M	mass	0.14 ± 0.02	9.3	< 0.0005	0.11 ± 0.02	4.5	< 0.0005
	β_{cl}	cl0024	-0.09 ± 0.02	3.9	< 0.0005	0.02 ± 0.03	0.5	0.6
	β_{ms}	ms0451	-0.11 ± 0.02	6.4	< 0.0005	-0.04 ± 0.03	1.4	0.2
3	β_0	constant				0.08 ± 0.02	4.2	< 0.0005
	β_M	mass				0.10 ± 0.02	4.4	< 0.0005
Is model 1 necesary over model 2?			likely		0.01	no		0.3
Is model 2 necessary over model 3?						no		0.2

The significance values (sig.) are the p-values from one-tailed t tests. The values will be twice for the two-tailed test. We use model 1 for [Fe/H] and model 3 for [Mg/H] relations with mass for further discussion.

redshift. This means that *normalizations* and *slopes* of the MZRs are allowed to depend on redshift. The model is

$$\begin{aligned} E([\text{Fe/H}] \mid M_{10}, \text{sample}) &= \beta_0 + \beta_M M_{10} + \beta_{cl} cl \\ &+ \beta_{ms} ms + \beta_{cl\cdot M}(cl \cdot M_{10}) + \beta_{ms\cdot M}(ms \cdot M_{10}) \end{aligned} \tag{5.1}$$

cl (or ms) is an indicator or a categorical variable with a value of 1 if the data point belongs to the Cl0024 (or MS0451) sample, and 0 otherwise. M_{10} is the galaxy stellar mass, defined as $M_{10} = \log[M_*/10^{10} M_\odot]$ so that the normalization is at $10^{10} M_\odot$. To digest this equation, we can view that β_0 is the normalization and β_M is the slope of the MZR for the $z \sim 0$ SDSS population; we call this the baseline MZR. The rest of the parameters are the evolution terms. β_{cl} is the change in the normalization, while $\beta_{cl\cdot M}$ is the change in the slope from those of local galaxies to those of $z \sim 0.4$ Cl0024 galaxies. β_{ms} and $\beta_{ms\cdot M}$ can be interpreted similarly.

We use the Monte Carlo (MC) technique with 1000 iterations to find the best-fit β_i parameters. In each iteration, we draw all samples randomly according to the measured probability distribution of [Fe/H]. We use the IDL code `regress` to find the best-fit parameters from each iteration. The final best-fit parameters (the mean and standard deviation) are listed as model 1 in Table 5.1.

Based on model 1, the MZRs of the higher redshift samples are likely steeper than that of local galaxies. The $\beta_{cl\cdot M}$ and $\beta_{ms\cdot M}$ terms suggest that the mean change in

[Fe/H] for a 1 dex increase in stellar mass is $\sim$ 0.9 dex more for higher redshift galaxies than for local galaxies. However, their significance values are only 2–3σ.

We further formally test the necessity of using the interaction model (model 1) over a simpler model (model 2) without the interactive terms, i.e., the model where the slopes are fixed to the same value regardless of redshift:

$$E([\mathrm{Fe/H}] \mid M_{10}, \mathrm{sample}) = \beta_0 + \beta_M M_{10} + \beta_{cl} cl + \beta_{ms} ms \tag{5.2}$$

The best-fit parameters are also listed in Table 5.1. The F-test for the R^2 change between the two models suggests that we can reject the null hypothesis that the simpler model 2 is sufficient to describe the data with a p-value of 0.01 (also $> 2\sigma$). Despite these borderline values of 2–3σ, for [Fe/H], we conclude that the interaction model (model 1) performs better than the simpler model 2 and we will work only with model 1 for further interpretations.

We found that the normalizations for the MZRs of the two higher redshift samples are statistically significantly lower ($> 3\sigma$) than the normalization of the local SDSS galaxies. With the more detailed spectral fitting model and larger sample size in this work, we confirm the finding in Chapter 4 that there is an evolution in the MZR when the metal indicator is [Fe/H]. The best estimate of normalizations at $10^{10} M_\odot$ for the $z \sim 0.39$ and the $z \sim 0.54$ samples are $\sim$ 0.14 and 0.18 $\pm$ 0.05 dex lower than the local samples. These values translate to the evolution of 0.04 $\pm$ 0.01 dex per observed Gyr, consistent with what we reported in Chapter 4.

In conclusion, we found that model 1 is the most appropriate to describe [Fe/H] evolution. We observe an evolution in [Fe/H] in both slope (likely at $2 - 3\sigma$) and normalization ($> 3\sigma$). This means that the amount of the evolution with observed redshift depends on mass. Over the same redshift range, lower mass quiescent galaxies evolve more strongly in [Fe/H] at fixed mass than higher mass quiescent galaxies. Further interpretation of [Fe/H] in Section 5.5 will be mainly based on the results from this model.

The MZR measured from [Mg/H] does not evolve with observed redshift

Now we look at the evolution of the relation between [Mg/H] and M_*. First, we repeat the fitting with model 1 but substitute [Fe/H] with [Mg/H]. We found that the mass–[Mg/H] relations of each redshift sample have the same slope within the $\sim 1\sigma$ uncertainties. The best-fit MZRs at different redshifts are roughly parallel. This suggests that the simpler model 2, in which slopes are fixed to a common value,

is more appropriate to describe the mass–[Mg/H] relation than the more complex model 1.

We continue to fit with model 2 and also found no significant evolution in the normalization of the MZR with observed redshift. As shown in Table 5.1, the values of β_{cl} and β_{ms} in model 2 are both consistent with zero. We found no significant evolution in [Mg/H] with observed redshift.

The results from the first two models suggest that we should proceed to fit a simple linear equation without redshift-dependent parameters (model 3). The parameters are also listed in Table 5.1. This model is most appropriate for the mass–[Mg/H] relation.

In summary, we did not detect an evolution of the stellar mass–[Mg/H] relation with redshift, neither in terms of slope nor normalization. This is in contrast to the $> 3\sigma$ detected in the evolution of the stellar mass–[Fe/H] relation with redshift. We can explicitly separate the best-fit parameters for both Fe and Mg abundances into the predicted MZRs at each redshift. The MZRs based on the measurements in this work are the following:

$$\begin{aligned}
\text{[Fe/H]}(M_{10}, z \sim 0) &= (-0.09 \pm 0.01) + (0.11 \pm 0.02)M_{10} \\
\text{[Fe/H]}(M_{10}, z \sim 0.39) &= (-0.23 \pm 0.05) + (0.20 \pm 0.06)M_{10} \\
\text{[Fe/H]}(M_{10}, z \sim 0.54) &= (-0.27 \pm 0.05) + (0.20 \pm 0.04)M_{10} \\
\text{[Mg/H]}(M_{10}) &= (+0.08 \pm 0.02) + (0.10 \pm 0.02)M_{10}
\end{aligned} \tag{5.3}$$

where $M_{10} = \log[M_*/10^{10}M_\odot]$.

5.5 Discussion

In this section, we will interpret the measured MZRs and their evolution with redshift, or lack thereof, using galactic chemical evolution models. Specifically, we use the measured MZRs to constrain average galactic outflows in terms of the mass-loading factor.

Comparison to galaxy simulations and semi-analytic models

The relatively shallow slopes of the MZRs estimated from our samples are consistent with other observations that specifically sample quiescent galaxies. Our slopes for both [Fe/H] and [Mg/H] are consistent within 2σ with the slopes derived from simple linear fits to the measurements by Choi, Conroy, et al. (2014), which were based on stacked spectra of field quiescent galaxies at similar redshifts. They are

also consistent with the slopes for the quiescent galaxies reported by Gallazzi, Bell, et al. (2014), in which metallicities (Z) were measured from a combination of H, Fe, and Mg absorption lines.

Based on measurements of the stellar metallicities of star-forming galaxies by Gallazzi, Bell, et al. (2014), the slopes of the MZRs of quiescent galaxies are shallower than those of star-forming galaxies.[2] Do we expect the stellar metallicities of quiescent galaxies to be different from those of star-forming galaxies?

Several semi-analytical (SA) and hydrodynamical simulation models have predicted stellar MZRs and their evolution with redshift (e.g. Lu, Wechsler, et al., 2014; Guo et al., 2016; Ma, Hopkins, Kasen, et al., 2016; Taylor et al., 2016). However, most results from simulations do not report metallicities based on galaxy star formation properties. The predictions are for the total populations, which are mixtures of star-forming and quiescent galaxies, depending on the passive fraction at each redshift.

It is therefore important to be cautious of interpreting the differences between observations and simulations. In Chapter 4, we compared the slope of [Fe/H] measured in quiescent galaxies to the slopes predicted in SA models in Lu, Wechsler, et al. (2014) at face value. We argued that the observational results agree the most with the model in which the outflow mass-loading factor from Type II SNe is independent of mass. However, the galaxies simulated by Lu, Wechsler, et al. have gas fractions on average ranging from 20% in the highest mass bin to 50% in the lowest mass bin at $z \sim 0$. In other words, portions of the simulated galaxies are still forming stars, which is not the case for our observed samples.

One study that specifically predicts the MZRs of quiescent galaxies is by Naiman et al. (2018), using the IllustrisTNG suite of simulations. The median of their simulated stellar abundances at $z \sim 0$ is plotted as magenta lines in Figure 5.3. We calculated the simulated [Mg/H] based on their median [Fe/H] and [Mg/Fe]. The slope in the simulated [Fe/H] is flatter than the observations. This was already explained by Naiman et al. as a result of the sampling criteria. Specifically, they excluded smaller mass galaxies in their sample. If those were included, they reported that the predicted mass–[Fe/H] relation could have been steeper. For [Mg/H], the simulated slope is $\sim$ 0.04 dex per log(mass), which is roughly half of what we observed. Because we calculated [Mg/H] from [Mg/Fe] and [Fe/H], it is likely that the cause of the difference in [Mg/H] slopes between the observations and simulations is the same as for [Fe/H].

[2] This is however not true for Local Group dwarf galaxies (Kirby, Cohen, et al., 2013).

Analytic chemical evolution model

As an alternative to more complex simulations, we can schematically understand the stellar metallicities of quiescent galaxies using an analytic galactic chemical evolution model. In this paper, we present a simple model that connects the metal mass in a quiescent galaxy to the average outflow it experienced. The model is based on the work of Lu, Blanc, et al. (2015). In this model, the assumptions are

1. Star formation occurs in the interstellar medium (ISM) that is perfectly mixed.
2. Metals are instantaneously recycled.
3. Outflows and inflows are permitted but the inflow do not contribute significantly to the total metal budget.[3]

We start by tracking the change in the total metal mass in the ISM ($dM_{z,g}$) as follows:

$$dM_{Z,g} = y dM_* - Z_g dM_* - \frac{\eta}{1-R} Z_g dM_* \tag{5.4}$$

where M_* is the mass of long-lived stars. The first term on the right side of equation 5.4 represents the metal mass that is newly produced and returned to the ISM by a generation of forming stars. y is the chemical yield, defined as the mass of metals returned to the ISM per mass turned into low-mass stars and remnants (dM_*). The second term is the metal mass in the ISM that is locked into stars. The last term is the metal mass that is lost to outflows, which is equal to the mass outflow rate times gas-phase metallicity Z_g. The outflow rate is parameterized in terms of the mass-loading factor (η), defined as the ratio of the mass outflow rate to the star formation rate (SFR). In this equation, the SFR is written in terms of dM_* and the return mass fraction R, which is the fraction of the mass of a stellar generation that returns—with its original composition—to the ISM from short-lived stars and stellar winds.

With the assumption that the ISM is well mixed, the change in the total metal mass locked in long-lived stars during the time interval in which a mass dM_* of long-lived stars formed is $dM_{Z,*} = Z_g dM_*$. We can substitute this to the last two terms of equation 5.4 to get

$$dM_{Z,g} = y dM_* - (1 + \frac{\eta}{1-R}) dM_{Z,*} \tag{5.5}$$

[3]We also derive an estimate for stellar metallicity when the inflow is enriched and evaluate its effect in Appendix 5.B.

We then integrate this equation, divide both sides by M_*, and rearrange the terms to get a description for the stellar metallicity of a galaxy:

$$Z_* = \frac{y - r_g Z_g}{1 + \frac{\langle \eta \rangle}{1-R}} \tag{5.6}$$

r_g is the cold gas fraction M_g/M_*. We adopt the definition of gas- and stellar-phase metallicity as $Z_g = M_{Z,g}/M_g$ and $Z_* = M_{Z,*}/M_*$. Therefore, in this equation, Z_g is the current gas-phase metallicity and Z_* is already an average metallicity, specifically the mass-weighted metallicity of the galaxy. Note that during the integration, we consider y and R as constant. $\langle \eta \rangle$ can be viewed as the average mass-loading factor weighted by $dM_{Z,*}$, which means that it is most sensitive to feedback at the galaxy's peak of star formation.

Now, we consider stellar metallicities of quiescent galaxies based on equation 5.6. We can see that at the limit where there is no gas left ($r_g = 0$), the second term in the numerator is zero and the stellar metallicity is a function of the mass-loading factor. This situation is appropriate to describe most quiescent galaxies.

We estimate that the product of gas fraction and gas-phase metallicity ($r_g Z_g$) in quiescent galaxies is at most 10% of the yield. First, quiescent galaxies have low gas fractions. In the local universe, early-type galaxies typically have molecular gas fractions less than a percent (Boselli et al., 2014). Galaxies at higher redshifts have higher molecular gas fractions. However, even at $z \sim 1.8$, an average early-type galaxy has a gas fraction that is less than 10% (Gobat et al., 2018). Second, quiescent galaxies do not have particularly high gas-phase metallicities. Griffith et al. (2019) measured gas-phase metallicities in 3 quiescent galaxies and found that their Z_g is roughly consistent with their stellar metallicities. This suggests that we can assume quiescent galaxies to have gas fractions r_g less than 10% and gas-phase metallicities less than the yield, y. Therefore, the $r_g Z_g$ term for quiescent galaxies in equation 5.6 is at least 1 dex smaller than the yield term.

Considering the discussion above, we estimate the stellar metallicities of quiescent galaxies to be

$$Z_{*,\text{quiescent}} \approx \frac{y}{1 + \frac{\langle \eta \rangle}{1-R}} \tag{5.7}$$

The yield and return fraction are fundamentally properties of stars. Though they can depend on IMF, they are often treated as constant and not a function of galaxy mass (e.g. Pagel, 1997). The only other variable in 5.7 is $\langle \eta \rangle$ (hereafter referred to

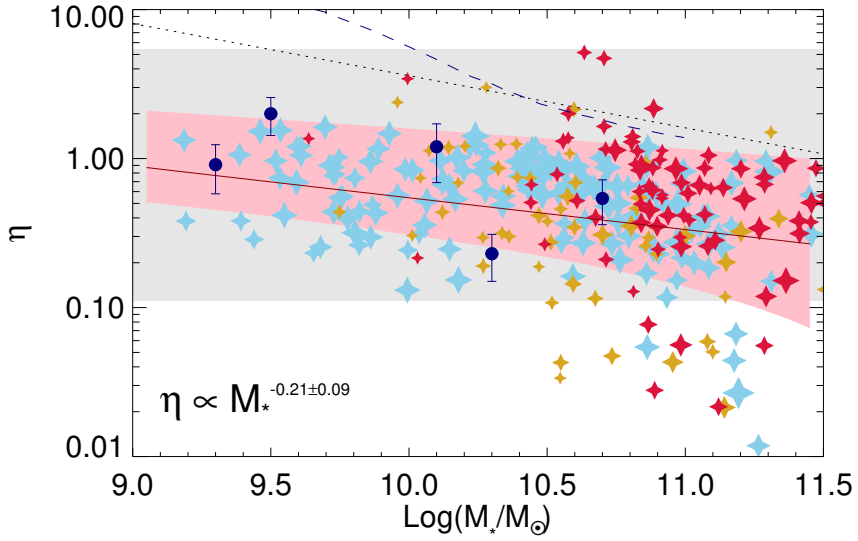

Figure 5.4: The relation between mass-loading factor (η) and galaxy stellar mass. Star-shaped symbols show η measured from our quiescent galaxy sample, color coded by redshift as in Figure 5.3. The median for the uncertainties in the estimated η is 0.1 dex. The line is the linear fit to individual measurements of η. Pink shading represents the best-fit and the scatter in the mass–[Mg/H] relation in Figure 5.3. The gray shading and navy dashed line are the estimations based on gas-phase metallicity from Spitoni et al. (2010) and Lu, Blanc, et al. (2015) (upper limits), respectively. Navy circles are the mass-loading factors measured from UV absorption lines (Chisholm et al., 2017). Results from the hydrodynamical simulation of Muratov et al. (2015) are shown as the dotted line.

simply as η). This means that a non-zero slope of the MZR of quiescent galaxies implies that the mass-loading factor is a function of galaxy mass.

Constraints on the mass-loading factor

We can now convert measured stellar metallicities of quiescent galaxies into mass-loading factors via equation 5.7. Because our model assumes instantaneous recycling, we apply our model to the observed Mg abundances. Mg is an α element produced in core-collapse supernovae and is appropriate for the assumption of instantaneous recycling. In addition, comparing to other α elements, it tracks oxygen the closest (Conroy, Graves, et al., 2014).

We estimate the mass ratio $M_{Mg,*}/M_*$ from the measured [Mg/H] using the solar abundance of Mg from Asplund et al. (2009) and the mass number of Mg (24) as the average Mg to H atomic weight in stars. The yield is set to 3 times the solar Mg abundance (Nomoto et al., 2006). Following Lu, Blanc, et al. (2015), we set the

return fraction to $R = 0.46$.

We found that mass-loading factor is a power-law function of galaxy mass. The result is plotted in Figure 5.4, where we convert each measurement of [Mg/H] to a mass-loading factor. We apply the MC technique to the probability distribution of [Mg/H] directly to obtain the best linear fit to the relation between $\log \eta$ and log mass. The mass-loading factor scales with galaxy stellar mass, $\eta \propto M_*^{-0.21\pm0.09}$. A galaxy that is 10 times less massive would have approximately 1.6 times larger outflow rate per SFR.[4]

We proposed an alternative way to constrain the mass-loading factor using the metallicities of quiescent galaxies. Our results agree reasonably well with other works that use chemical evolution models for star-forming galaxies. In Figure 5.4, we plot as a blue dashed line the constraints from Lu, Blanc, et al. (2015), whose model is very similar to ours (and actually was the model that inspired our work). Their results are upper limits because of the uncertainties in gas fractions, different gas-phase metallicity calibrations, and a conservative choice of yield (5 times solar). Our estimates of η are consistent with the upper limit. We also plot an earlier estimate from Spitoni et al. (2010) (in gray shading). Their estimate is based on a chemical evolution model of the gas-phase MZR that allows both infall and outflow.

Our estimates for the mass-loading factors are consistent with other observational methods. A direct way to constrain outflow is to observe absorption features by the outflow of material seen in galaxy or background quasar spectra. Alternatively, some emission lines from the outflow gas can be observed. Although these methods are able to measure the current outflow rate directly, they are still limited in terms of sample size. The results from applying these techniques to local star-forming galaxies suggest a range in mass-loading factor from 0.1 to several tens (e.g., Bolatto et al., 2013; Bouché et al., 2012). In figure 5.4, we plot in navy circles the results based on UV absorption lines from Chisholm et al. (2017), in which stellar mass measurements were readily available. Those results agree with ours not only in terms of slope but also normalization. In other words, our archaeologically derived mass-loading factors are consistent with those that are measured directly in currently star-forming galaxies.

[4]In fitting this linear relation between $\log \eta$ and log mass, we exclude 18 galaxies whose best estimated [Mg/H]'s are larger than the adopted yield. If we instead measured η with an extremely high yield (6 times the solar abundance of Mg), the resulting power-law index would be lower ($\eta = -0.13 \pm 0.05$) but still consistent with the current estimate. With this assumption of yield, the normalization would be ~ 0.5 dex higher.

Given the simplicity of our chemical evolution model, the estimated power-law index of the relation between η and M_* agrees surprisingly well with both analytic and hydrodynamical simulations. Hayward et al. (2017) presented an analytic model for galactic outflows driven by stellar feedback. In the model, the SFR is self-regulated. It is set by an equilibrium between the momentum injection rate from stellar feedback and the dissipation rate by turbulence and occasional outflows. The outflow occurs when the momentum accumulated within a coherence time is large enough given the gas density of the ISM patch. Their model predicts that the mass-loading factor scales with gas fraction and stellar mass. For the stellar mass component, their predicted scaling relation is $\eta \propto M_*^{-0.23}$, which is consistent with our observation.

However, the observed exponent of the scaling relation is smaller than predicted in an earlier semi-analytical model by Lagos et al. (2013). In this model, the basic assumptions are similar to those of Hayward et al.: the ISM disk is supported by turbulent pressure, and stellar feedback drives gas outflow. However, the timescale of the Lagos et al. simulation is set by the lifetime of giant molecular clouds, which can be larger than the crossing time adopted by Hayward et al., and their SFR is set by an empirical relation. They also found that the mass-loading factor scales with gas density and galaxy mass with an exponent of $\sim -1.1 \pm 0.5$ (based on estimation from their Figure 14), which is larger than what we observed.

Hydrodynamical simulations predict scaling relations of the mass-loading factor that are consistent in slope but with somewhat larger normalization than the observational result. The dotted line in figure 5.4 shows the scaling relation presented by Muratov et al. (2015) based on the Feedback in Realistic Environments (FIRE) simulations. The slope is consistent with our observation but it is $\sim 3 - 7$ times larger in normalization than our best-fit line.

It is normally difficult for simulations to define mass-loading factors that match observations. Muratov et al. (2015) calculated the mass-loading factors from outflow flux (all gas with outward radial velocity) at $0.25R_{vir}$. This does not capture the effect of recycling from galactic fountains, whereas our measurement does. Thus, their results are likely to be larger than our measurements by construction. Based on the IllustrisTNG results from Pillepich et al. (2018), in which their definition of the mass-loading factor was explicitly remarked as not applicable to observations and was calculated from wind energy at injection, their "mass-loading factor" scales with halo mass M_{200} with the power law index of ~ -0.8 to -0.6. If we convert halo mass to stellar mass using the stellar-to-halo mass relation $M_* \propto M_{200}^{\sim 2.27}$, found in the

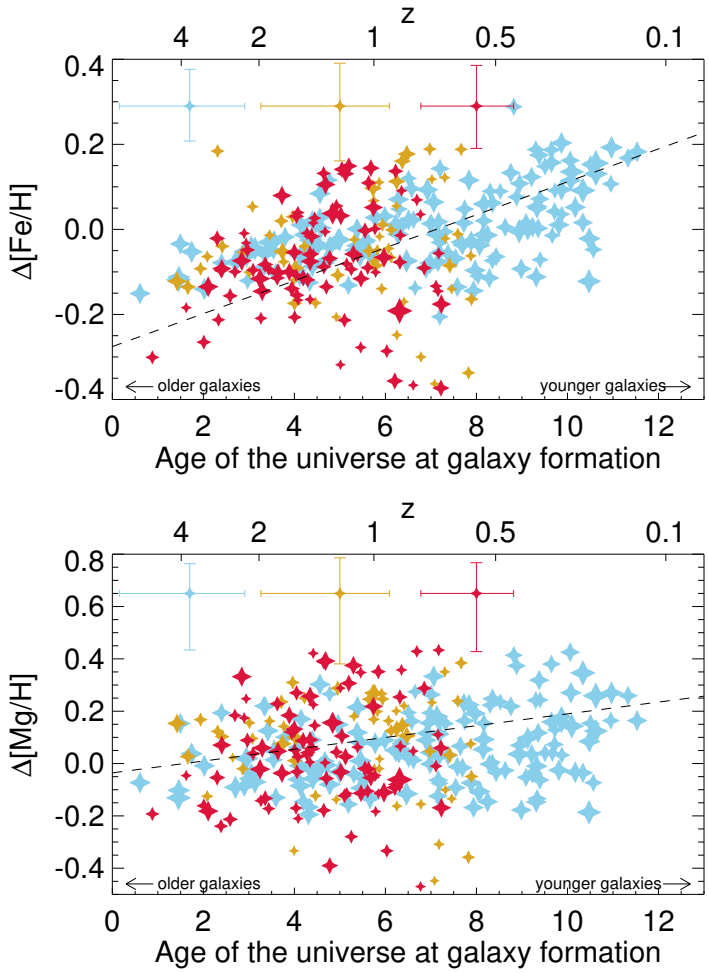

Figure 5.5: The dependence of [Fe/H] and [Mg/H] on formation redshift when the dependence on mass is removed (see Section 5.5). The median uncertainties are shown in the top of each panel. The color scheme follows previous plots. The slope in the top panel is statistically significant, but the slope in the bottom panel is not.

same suite of simulations (Niemiec et al., 2018) for galaxies with $M_* < 10^{11.2} M_\odot$, the power-law index with stellar mass is approximately -0.35 to -0.26, which is still roughly consistent with what we measured.

Remarks on the MZR evolution with redshift

In Chapter 4, we argued that the evolution of metallicity with formation redshift in quiescent galaxies is more fundamental than the evolution with observed redshift. Galaxies of the same mass that formed at the same redshift should have similar metallicities regardless of when we observe them, assuming that most quiescent

galaxies evolve passively.[5] As in Chapter 4, we trace each galaxy back to its formation time by adding the age of the universe when it is observed to the age of the galaxy. We plot the deviation in metallicities from the $z \sim 0$ MZR relation as a function of formation time in Figure 5.5. The plots essentially show the dependence of metallicities on formation redshift (age), while the dependence on mass is removed.

The deviation in the [Fe/H] abundances is consistent with what we found in Chapter 4. The [Fe/H] evolution with formation time is 0.04 ± 0.01 dex per formation Gyr. The average difference between the formation times of the $z \sim 0.39$ (or $z \sim 0.54$) and the local sample is 1.7 (or 2.2 Gyr). Multiplying these age differences with the amount of evolution with formation time, we recover the magnitude of evolution with observed redshift that is consistent with the results in Table 5.1.

However, we do not find a significant dependency of [Mg/H] on formation time. As shown in the right of Figure 5.5, the best-fit slope for the evolution is 0.02 ± 0.01 dex per formation Gyr, a significance less than 2σ. An evolution of [Mg/H] with formation time would imply that fundamental properties of galaxies such as yield or mass-loading factors depend on time. Since we did not detect any significant evolution of [Mg/H] with both observed and formation redshift, these fundamental properties are not required to depend on redshift.

The absence of the evolution of [Mg/H] with redshift together with the existence of the evolution of [Fe/H] with redshift suggest that the latter is caused by the delayed time of the Type Ia supernovae. The main difference between the two elements is that Mg is approximately instantaneously recycled while the Fe can have a delay time up to several Gyrs (e.g., Maoz et al., 2012). Thus, the evolution of [Fe/H] with redshift in quiescent galaxies that we found is not surprising and is perhaps expected due to selection effects. We restricted our sample to be passive. The galaxies at higher redshift are required to finish forming star early, while the galaxies at lower redshift could have quenched at later time. Quiescent galaxies at higher redshift had less time to become enriched by Fe from Type Ia supernovae.

The evolution of [Fe/H] with redshift we found in this work does not necessarily contradict the non-detection of evolution of MZR with redshift in recent literature, mainly due to large measurement uncertainties. We first note that most of these

[5] If the sample size were large enough, we could also use this technique to trace whether quiescent galaxies evolve passively in metal abundance. This could be done by comparing MZRs of galaxies that formed at the same formation redshift but are observed at different redshifts.

works measure total metallicity [Z/H] without α enhancement. However, most of the spectra either do not cover the Mg b lines, or the absorption indices were chosen to be less sensitive to α enhancement. Therefore these [Z/H] should be close to our [Fe/H] measurements. Regardless, Gallazzi, Bell, et al. (2014), Estrada-Carpenter et al. (2019) and Saracco et al. (2019) found that the [Z/H]–mass relations in quiescent galaxies (both in the field and in clusters) at $z \sim 1.1 - 1.6$ are not different from that of local galaxies. These studies are quite limited in terms of sample size, especially for the lower mass galaxies, and either small wavelength coverage or coarse spectral resolution. These aspects result in large measurement uncertainties that can disguise the expected evolution.

For example, based on the mean age of 2.0 ± 0.8 Gyr of 7 quiescent galaxies at $z \sim 1.22$ in Saracco et al.'s (2019) sample and the amount of evolution with formation time found in this work, we expect them to have an average 0.15 ± 0.06 lower metallicity than the local galaxies. This number is smaller than their mean uncertainty in [Z/H] measurements (0.22 dex), standard deviation within the sample (0.18 dex), and uncertainty in the normalization of the best-fit linear relation with stellar mass (0.15 dex).

Lastly, we do not expect our finding to be a result of environment as already discussed in Chapter 4. In short, a cluster environment mainly affects the quenched fraction of galaxies in the sense that cluster galaxies tend to be more quenched than field galaxies. However, in terms of chemical composition, the effect is either small (< 0.06 dex, e.g., Cooper, Tremonti, et al., 2008; Jørgensen et al., 2018), or not significant (e.g., Harrison et al., 2011; Kacprzak et al., 2015; Fitzpatrick et al., 2015)

5.6 Conclusion

We measured ages and metallicities ([Fe/H] and [Mg/H]) of individual quiescent galaxies in two galaxy clusters at $z \sim 0.39$ and $z \sim 0.54$ using full-spectrum fitting. The sample consists of 62 galaxies at $z = 0.39$ and 92 galaxies at $z = 0.54$ in the galaxy clusters Cl0024 and MS0451. We used the SPS models and response functions from Conroy, Gunn, and M. White (2009) and Conroy, Villaume, et al. (2018). By expanding the sample size from Leethochawalit, Kirby, et al. (2018) and measuring Mg abundances in addition to Fe, we were able archaeologically to derive the dependence of mass-loading factor on mass based on the mass–metallicity relation of quiescent galaxies. The summary of the findings this paper are the following.

- We confirmed the finding in Chapter 4 that the stellar mass–[Fe/H] relation evolves with redshift. At a fixed stellar mass, quiescent galaxies at higher redshift have lower [Fe/H] than quiescent galaxies at lower redshift. The slope of the relation is also likely steeper at higher redshift.

- We found no evolution in the mass–[Mg/H] relation with redshift. This suggests that the evolution observed in the [Fe/H] abundance is due to the delay time of Type Ia supernovae.

- We constrained the mass-loading factor using an analytic chemical evolution model for quiescent galaxies, assuming that Mg is an indicator of instantaneously recycled elements. We found that the mass-loading factor is a power-law function of galaxy stellar mass, $\eta \propto M_*^{-0.21\pm0.09}$, over the observed mass range $\sim 10^{9.5}$ to $10^{11.5} M_\odot$.

- Our constraint on the mass-loading factor is consistent with an analytic prediction in which outflow is caused by star-formation feedback in a turbulent disk (Hayward et al., 2017). It is also consistent with the results from direct measurements of outflow using UV absorption lines.

There are still many open questions that can be explored beyond this paper. These include confirming the stellar mass–metallicity relationship (in both [Fe/H] and alpha elements) with a large sample of field galaxies, especially down to low masses ($\sim 10^{9.5}\ M_\odot$). As suggested in Section 5.5, a large sample of abundance measurements in field quiescent galaxies at different redshifts would allow us not only to even more firmly establish the evolution of MZRs, but also to determine chemically the degree of passive evolution of quiescent galaxies. Beyond this, precise and consistent measurements of stellar metallicities in star-forming galaxies (especially as a function of gas fraction), will help bridge and strengthen the findings in our work with those by Gallazzi, Charlot, Brinchmann, S. D. M. White, and Tremonti (2005) and Lu, Blanc, et al. (2015).

The authors acknowledge Charlie Conroy, Philip Hopkins, and Gwen Rudie for useful feedback and the use of their models and fitting codes.

Table 5.2: Catalog of Measured Age and Metalicities of Cl0024+17 galaxies

No.	RA	DEC	$\log(M_*/M_\odot)$	Age (Gyr)	[Fe/H]	[Mg/Fe]
1	00 25 51.07	+17 08 42.4	10.8	$2.2^{2.6}_{2.0}$	$+0.19^{+0.28}_{+0.06}$	$+0.02^{+0.17}_{-0.11}$

Continued on next page

Table 5.2 – *Continued from previous page*

No.	RA	DEC	$\log(M_*/M_\odot)$	Age (Gyr)	[Fe/H]	[Mg/Fe]
2	00 25 54.52	+17 16 26.4	10.4	$3.0^{4.9}_{1.9}$	$-0.30^{-0.06}_{-0.46}$	$+0.25^{+0.45}_{+0.03}$
3	00 25 57.73	+17 08 01.5	10.4	$3.5^{5.9}_{2.5}$	$+0.06^{+0.18}_{-0.03}$	$+0.04^{+0.20}_{-0.14}$
4	00 26 04.30	+17 18 46.5	10.3	$2.2^{3.1}_{1.5}$	$-0.42^{-0.19}_{-0.56}$	$+0.08^{+0.40}_{-0.22}$
5	00 26 04.44	+17 20 00.6	10.5	$2.9^{4.5}_{0.5}$	$-0.21^{+0.05}_{-0.35}$	$+0.56^{+0.77}_{+0.34}$
6	00 26 05.80	+17 19 19.0	10.9	$7.2^{9.7}_{5.5}$	$-0.06^{+0.05}_{-0.14}$	$+0.10^{+0.22}_{-0.04}$
7	00 26 05.84	+17 19 19.0	10.9	$5.5^{7.6}_{3.6}$	$-0.07^{+0.02}_{-0.15}$	$+0.38^{+0.51}_{+0.25}$
8	00 26 06.95	+17 19 42.8	10.8	$5.2^{6.1}_{3.7}$	$-0.09^{-0.03}_{-0.15}$	$+0.35^{+0.44}_{+0.27}$
9	00 26 09.66	+17 11 13.5	10.4	$9.3^{12.1}_{7.6}$	$-0.44^{-0.32}_{-0.55}$	$+0.59^{+0.72}_{+0.45}$
10	00 26 15.16	+17 18 15.6	10.4	$3.5^{5.7}_{2.7}$	$-0.12^{+0.05}_{-0.33}$	$+0.22^{+0.38}_{+0.05}$
11	00 26 18.45	+17 07 01.1	10.3	$3.9^{6.4}_{2.1}$	$-0.09^{+0.01}_{-0.30}$	$+0.37^{+0.52}_{+0.23}$
12	00 26 21.49	+17 14 11.8	10.6	$6.4^{8.2}_{4.5}$	$-0.08^{-0.01}_{-0.16}$	$+0.20^{+0.32}_{+0.07}$
13	00 26 22.90	+17 12 31.4	11.1	$2.8^{3.7}_{2.9}$	$+0.20^{+0.29}_{+0.11}$	$+0.14^{+0.22}_{+0.02}$
14	00 26 22.91	+17 12 31.3	11.1	$3.4^{5.1}_{2.6}$	$-0.01^{+0.13}_{-0.09}$	$+0.40^{+0.56}_{+0.25}$
15	00 26 24.82	+17 12 21.5	10.9	$6.2^{7.8}_{4.5}$	$-0.09^{-0.01}_{-0.21}$	$+0.42^{+0.54}_{+0.30}$
16	00 26 25.54	+17 12 09.0	10.6	$4.3^{5.8}_{0.1}$	$-0.23^{+0.07}_{-0.40}$	$+0.23^{+0.44}_{+0.02}$
17	00 26 27.12	+17 12 25.9	10.9	$6.1^{7.8}_{4.9}$	$+0.06^{+0.16}_{-0.02}$	$+0.20^{+0.31}_{+0.09}$
18	00 26 27.98	+17 11 37.7	10.5	$9.2^{11.2}_{6.3}$	$-0.15^{-0.04}_{-0.24}$	$+0.60^{+0.74}_{+0.45}$
19	00 26 28.40	+17 11 16.6	10.6	$1.9^{2.6}_{1.5}$	$+0.10^{+0.26}_{-0.11}$	$+0.39^{+0.58}_{+0.20}$
20	00 26 29.11	+17 10 24.7	10.7	$4.3^{5.7}_{3.1}$	$-0.12^{-0.04}_{-0.25}$	$+0.56^{+0.68}_{+0.44}$
21	00 26 29.50	+17 10 32.6	10.5	$3.3^{7.2}_{1.9}$	$-0.17^{+0.02}_{-0.40}$	$+0.57^{+0.75}_{+0.42}$
22	00 26 29.92	+17 10 06.8	10.3	$1.4^{6.0}_{1.1}$	$-0.22^{-0.14}_{-0.54}$	$+0.28^{+0.51}_{+0.03}$
23	00 26 30.76	+17 12 26.3	10.3	$1.6^{2.1}_{1.4}$	$-0.07^{-0.02}_{-0.15}$	$+0.42^{+0.61}_{+0.22}$
24	00 26 31.04	+17 11 09.3	10.4	$3.5^{5.5}_{2.5}$	$+0.02^{+0.14}_{-0.06}$	$+0.27^{+0.41}_{+0.13}$
25	00 26 31.22	+17 12 08.4	10.7	$3.0^{3.7}_{2.6}$	$+0.10^{+0.25}_{-0.01}$	$+0.18^{+0.29}_{+0.07}$
26	00 26 31.40	+17 17 00.3	11.1	$5.0^{7.6}_{3.7}$	$-0.01^{+0.09}_{-0.07}$	$+0.45^{+0.56}_{+0.33}$
27	00 26 31.41	+17 10 55.6	11.3	$7.6^{9.4}_{5.6}$	$-0.08^{-0.00}_{-0.15}$	$+0.31^{+0.41}_{+0.23}$
28	00 26 31.41	+17 10 27.1	10.0	$5.2^{10.1}_{3.0}$	$-0.22^{-0.04}_{-0.51}$	$-0.04^{+0.14}_{-0.20}$
29	00 26 31.53	+17 06 03.7	11.3	$2.0^{2.4}_{1.5}$	$-0.36^{-0.10}_{-0.58}$	$+0.26^{+0.54}_{-0.01}$
30	00 26 31.56	+17 17 12.8	10.5	$6.8^{8.4}_{4.6}$	$-0.08^{-0.01}_{-0.16}$	$+0.31^{+0.42}_{+0.21}$
31	00 26 31.70	+17 11 25.4	10.0	$3.3^{8.3}_{1.4}$	$-0.10^{+0.06}_{-0.43}$	$+0.21^{+0.40}_{-0.00}$
32	00 26 32.50	+17 10 26.0	10.9	$3.1^{5.0}_{1.8}$	$-0.06^{+0.14}_{-0.22}$	$+0.35^{+0.50}_{+0.20}$
33	00 26 32.71	+17 07 56.0	10.3	$2.5^{3.5}_{1.7}$	$-0.35^{-0.10}_{-0.55}$	$+0.31^{+0.55}_{+0.06}$
34	00 26 33.32	+17 17 26.9	10.2	$6.6^{7.2}_{5.4}$	$-1.08^{-0.99}_{-1.14}$	$+1.20^{+1.33}_{+0.99}$
35	00 26 33.54	+17 09 23.9	10.6	$1.8^{2.2}_{1.6}$	$-0.11^{-0.05}_{-0.28}$	$+0.28^{+0.44}_{+0.12}$
36	00 26 33.60	+17 09 20.2	10.5	$9.2^{11.4}_{6.5}$	$-0.06^{+0.06}_{-0.15}$	$+0.31^{+0.43}_{+0.19}$
37	00 26 33.81	+17 12 16.6	10.6	$2.2^{5.2}_{1.4}$	$+0.09^{+0.36}_{-0.29}$	$+0.28^{+0.48}_{+0.06}$
38	00 26 34.35	+17 10 22.1	11.1	$3.5^{4.6}_{2.9}$	$-0.05^{+0.06}_{-0.10}$	$+0.51^{+0.61}_{+0.41}$
39	00 26 34.59	+17 10 16.4	10.8	$5.2^{7.5}_{4.0}$	$-0.17^{-0.09}_{-0.38}$	$+0.35^{+0.46}_{+0.24}$
40	00 26 34.98	+17 10 21.3	10.5	$5.3^{7.7}_{4.0}$	$+0.00^{+0.09}_{-0.06}$	$+0.44^{+0.55}_{+0.34}$
41	00 26 35.70	+17 09 43.1	11.5	$7.8^{9.4}_{5.8}$	$-0.04^{+0.03}_{-0.11}$	$+0.42^{+0.51}_{+0.35}$
42	00 26 37.27	+17 10 00.2	10.6	$6.9^{8.3}_{6.2}$	$+0.16^{+0.23}_{+0.07}$	$+0.10^{+0.20}_{-0.00}$

Continued on next page

Table 5.2 – *Continued from previous page*

No.	RA	DEC	$\log(M_*/M_\odot)$	Age (Gyr)	[Fe/H]	[Mg/Fe]
43	00 26 37.52	+17 09 08.7	10.6	$5.5^{7.4}_{4.5}$	$-0.00^{+0.08}_{-0.06}$	$+0.38^{+0.48}_{+0.28}$
44	00 26 37.90	+17 09 22.0	11.2	$5.5^{7.4}_{4.5}$	$-0.01^{+0.08}_{-0.05}$	$+0.28^{+0.36}_{+0.20}$
45	00 26 37.91	+17 09 37.8	10.7	$3.3^{4.8}_{2.6}$	$-0.00^{+0.11}_{-0.07}$	$+0.40^{+0.52}_{+0.28}$
46	00 26 38.41	+17 09 58.7	10.3	$3.1^{7.5}_{1.4}$	$-0.10^{+0.08}_{-0.36}$	$+0.38^{+0.54}_{+0.21}$
47	00 26 38.65	+17 09 14.5	9.7	$5.8^{7.4}_{3.8}$	$-0.13^{-0.05}_{-0.24}$	$+0.35^{+0.48}_{+0.23}$
48	00 26 38.80	+17 09 59.5	10.7	$1.6^{1.9}_{1.4}$	$+0.17^{+0.23}_{+0.11}$	$+0.36^{+0.51}_{+0.21}$
49	00 26 39.99	+17 09 41.8	10.1	$1.9^{2.9}_{1.4}$	$-0.10^{+0.13}_{-0.42}$	$+0.09^{+0.36}_{-0.18}$
50	00 26 41.16	+17 10 01.8	11.1	$3.6^{5.4}_{2.6}$	$-0.02^{+0.09}_{-0.09}$	$+0.45^{+0.57}_{+0.34}$
51	00 26 43.23	+17 08 41.0	10.1	$3.1^{6.8}_{1.9}$	$-0.08^{+0.05}_{-0.38}$	$+0.30^{+0.45}_{+0.16}$
52	00 26 43.70	+17 07 12.8	10.0	$3.0^{8.3}_{1.2}$	$-0.04^{+0.17}_{-0.43}$	$+0.32^{+0.51}_{+0.13}$
53	00 26 48.22	+17 10 46.3	10.2	$1.8^{2.6}_{1.3}$	$-0.21^{+0.02}_{-0.54}$	$+0.18^{+0.47}_{-0.13}$
54	00 26 48.30	+17 12 35.4	10.6	$1.4^{6.2}_{1.2}$	$-0.36^{-0.29}_{-0.49}$	$+0.14^{+0.35}_{-0.06}$
55	00 26 51.84	+17 08 39.9	10.5	$7.3^{8.9}_{5.7}$	$-0.13^{-0.03}_{-0.23}$	$+0.42^{+0.53}_{+0.33}$
56	00 26 54.38	+17 08 27.0	11.0	$3.5^{4.5}_{3.0}$	$+0.11^{+0.20}_{+0.04}$	$+0.33^{+0.42}_{+0.24}$
57	00 27 07.76	+17 10 49.5	11.0	$2.7^{4.3}_{2.6}$	$+0.20^{+0.30}_{+0.03}$	$+0.09^{+0.19}_{-0.03}$
58	00 27 09.82	+17 11 19.3	10.5	$1.4^{6.1}_{1.1}$	$-0.09^{-0.02}_{-0.51}$	$+0.30^{+0.59}_{+0.02}$
59	00 27 25.16	+17 07 22.4	10.1	$4.8^{8.5}_{3.2}$	$-0.25^{-0.05}_{-0.43}$	$+0.22^{+0.36}_{+0.07}$

Table 5.3: Catalog of Measured Age and Metalicities of MS0451 galaxies

No.	RA	DEC	$\log(M_*/M_\odot)$	Age (Gyr)	[Fe/H]	[Mg/Fe]
1	04 53 18.17	-02 58 57.7	11.1	$3.4^{6.0}_{2.4}$	$+0.05^{+0.18}_{-0.04}$	$+0.56^{+0.72}_{+0.41}$
2	04 53 33.42	-02 56 23.5	10.9	$1.9^{2.2}_{1.6}$	$-0.07^{-0.01}_{-0.19}$	$+0.49^{+0.65}_{+0.34}$
3	04 53 36.54	-03 04 13.5	11.3	$4.8^{6.8}_{3.4}$	$-0.06^{+0.03}_{-0.12}$	$+0.50^{+0.61}_{+0.39}$
4	04 53 38.64	-02 54 11.2	10.8	$3.0^{4.4}_{2.3}$	$-0.08^{+0.13}_{-0.21}$	$+0.13^{+0.28}_{-0.05}$
5	04 53 52.61	-02 55 31.5	10.8	$1.9^{2.3}_{1.6}$	$-0.35^{-0.27}_{-0.58}$	$+0.42^{+0.62}_{+0.26}$
6	04 53 54.23	-03 04 53.4	11.1	$2.4^{3.0}_{2.1}$	$-0.25^{-0.09}_{-0.39}$	$+0.78^{+0.89}_{+0.64}$
7	04 53 57.58	-03 01 18.7	10.6	$2.2^{2.7}_{1.8}$	$-0.31^{-0.09}_{-0.49}$	$+0.12^{+0.33}_{-0.09}$
8	04 53 58.37	-03 06 51.6	10.8	$1.5^{1.9}_{1.3}$	$+0.07^{+0.12}_{-0.08}$	$+0.52^{+0.70}_{+0.33}$
9	04 53 59.17	-02 56 51.1	11.5	$3.1^{3.9}_{2.9}$	$+0.10^{+0.22}_{-0.04}$	$+0.09^{+0.19}_{-0.03}$
10	04 54 00.29	-02 58 58.7	10.2	$3.5^{4.8}_{2.8}$	$+0.04^{+0.13}_{-0.02}$	$+0.46^{+0.55}_{+0.37}$
11	04 54 02.29	-02 57 09.8	11.0	$2.2^{2.4}_{1.9}$	$-0.05^{+0.05}_{-0.13}$	$+0.11^{+0.24}_{-0.03}$
12	04 54 02.34	-03 00 47.9	10.5	$2.4^{3.0}_{2.0}$	$-0.03^{+0.12}_{-0.19}$	$+0.12^{+0.28}_{-0.08}$
13	04 54 02.70	-02 56 34.8	11.2	$4.9^{6.1}_{3.6}$	$-0.07^{-0.02}_{-0.12}$	$+0.25^{+0.33}_{+0.17}$
14	04 54 02.92	-02 59 03.5	10.5	$2.3^{2.5}_{2.0}$	$-0.50^{-0.41}_{-0.63}$	$+0.98^{+1.03}_{+0.89}$
15	04 54 03.11	-02 55 56.1	11.5	$4.0^{5.5}_{3.2}$	$+0.11^{+0.20}_{+0.02}$	$+0.14^{+0.25}_{+0.01}$
16	04 54 03.62	-03 03 38.5	10.7	$5.5^{6.9}_{3.6}$	$-0.11^{-0.03}_{-0.20}$	$+0.45^{+0.56}_{+0.34}$
17	04 54 04.41	-03 01 42.2	11.1	$3.1^{4.6}_{2.3}$	$-0.18^{-0.04}_{-0.40}$	$+0.25^{+0.38}_{+0.11}$
18	04 54 04.91	-03 00 52.6	10.5	$5.3^{7.9}_{4.1}$	$-0.05^{+0.05}_{-0.12}$	$+0.35^{+0.49}_{+0.21}$

Continued on next page

Table 5.3 – *Continued from previous page*

No.	RA	DEC	$\log(M_*/M_\odot)$	Age (Gyr)	[Fe/H]	[Mg/Fe]
19	04 54 04.94	-03 05 45.6	11.4	$1.1^{1.2}_{1.0}$	$-0.31^{-0.28}_{-0.33}$	$+0.58^{+0.80}_{+0.36}$
20	04 54 05.41	-02 57 19.7	10.4	$3.5^{5.5}_{2.4}$	$-0.05^{+0.11}_{-0.18}$	$+0.17^{+0.34}_{-0.01}$
21	04 54 05.73	-03 00 40.9	11.1	$5.9^{7.1}_{4.8}$	$+0.02^{+0.10}_{-0.04}$	$+0.12^{+0.21}_{+0.02}$
22	04 54 06.52	-02 59 28.6	11.0	$4.1^{5.6}_{3.4}$	$-0.03^{+0.03}_{-0.10}$	$+0.34^{+0.44}_{+0.25}$
23	04 54 06.54	-02 56 33.7	10.9	$3.4^{4.5}_{2.7}$	$-0.06^{+0.08}_{-0.12}$	$-0.16^{-0.04}_{-0.28}$
24	04 54 06.89	-02 59 04.2	10.8	$4.7^{6.3}_{3.1}$	$-0.09^{-0.01}_{-0.21}$	$+0.08^{+0.21}_{-0.06}$
25	04 54 06.92	-03 02 08.6	10.8	$2.6^{3.7}_{2.2}$	$-0.11^{+0.02}_{-0.31}$	$+0.16^{+0.29}_{+0.02}$
26	04 54 07.04	-02 59 45.2	11.0	$3.9^{5.5}_{3.0}$	$-0.06^{+0.02}_{-0.12}$	$+0.49^{+0.59}_{+0.40}$
27	04 54 07.28	-02 59 40.2	11.1	$6.6^{8.9}_{4.6}$	$-0.15^{-0.04}_{-0.33}$	$+0.29^{+0.43}_{+0.14}$
28	04 54 07.39	-03 05 30.1	10.9	$3.7^{5.1}_{2.9}$	$-0.00^{+0.11}_{-0.07}$	$+0.23^{+0.33}_{+0.13}$
29	04 54 07.53	-02 57 48.5	10.6	$4.1^{8.4}_{2.6}$	$-0.19^{+0.03}_{-0.51}$	$+0.12^{+0.30}_{-0.07}$
30	04 54 07.91	-03 03 01.1	10.7	$5.8^{8.6}_{4.4}$	$-0.22^{-0.09}_{-0.35}$	$-0.29^{-0.18}_{-0.35}$
31	04 54 08.34	-02 59 15.8	10.8	$5.7^{7.2}_{3.8}$	$-0.10^{-0.03}_{-0.18}$	$+0.02^{+0.14}_{-0.11}$
32	04 54 08.49	-03 01 08.4	11.0	$2.9^{3.4}_{2.4}$	$-0.04^{+0.09}_{-0.14}$	$+0.59^{+0.70}_{+0.48}$
33	04 54 08.49	-03 00 33.7	10.9	$5.4^{6.8}_{3.7}$	$-0.07^{-0.00}_{-0.13}$	$+0.57^{+0.66}_{+0.47}$
34	04 54 08.54	-03 00 31.3	10.9	$8.2^{10.3}_{6.5}$	$+0.04^{+0.15}_{-0.05}$	$+0.12^{+0.24}_{-0.01}$
35	04 54 08.70	-03 00 54.7	10.8	$3.8^{6.2}_{2.4}$	$-0.01^{+0.08}_{-0.08}$	$+0.20^{+0.32}_{+0.07}$
36	04 54 08.98	-03 01 57.8	10.8	$3.8^{5.7}_{2.4}$	$-0.06^{+0.08}_{-0.16}$	$+0.44^{+0.60}_{+0.28}$
37	04 54 09.25	-03 00 50.7	11.3	$8.2^{10.1}_{6.6}$	$-0.03^{+0.05}_{-0.11}$	$+0.13^{+0.21}_{+0.04}$
38	04 54 09.36	-02 57 41.6	11.6	$2.4^{2.8}_{2.2}$	$+0.14^{+0.23}_{+0.01}$	$+0.32^{+0.42}_{+0.22}$
39	04 54 09.62	-03 01 00.9	11.3	$4.8^{6.0}_{3.7}$	$-0.09^{-0.05}_{-0.15}$	$+0.36^{+0.45}_{+0.27}$
40	04 54 09.74	-03 00 15.5	11.0	$2.7^{4.1}_{2.2}$	$-0.04^{+0.13}_{-0.24}$	$+0.28^{+0.45}_{+0.11}$
41	04 54 09.88	-02 58 44.9	10.9	$4.9^{9.7}_{3.4}$	$-0.20^{-0.08}_{-0.49}$	$+0.23^{+0.37}_{+0.08}$
42	04 54 09.93	-03 01 29.8	10.7	$5.2^{6.8}_{3.6}$	$-0.06^{+0.01}_{-0.13}$	$+0.30^{+0.41}_{+0.20}$
43	04 54 10.08	-03 00 50.4	10.9	$3.7^{5.7}_{2.8}$	$+0.04^{+0.14}_{-0.04}$	$+0.28^{+0.40}_{+0.15}$
44	04 54 10.32	-02 59 15.8	10.8	$5.6^{7.5}_{4.4}$	$-0.16^{-0.07}_{-0.32}$	$+0.10^{+0.21}_{-0.01}$
45	04 54 10.55	-02 59 32.3	11.1	$4.1^{5.5}_{3.0}$	$+0.07^{+0.15}_{-0.01}$	$+0.39^{+0.49}_{+0.30}$
46	04 54 10.88	-03 01 36.6	11.2	$3.0^{4.1}_{2.8}$	$+0.20^{+0.30}_{+0.07}$	$+0.31^{+0.42}_{+0.21}$
47	04 54 11.17	-02 59 37.3	10.9	$4.2^{8.9}_{2.8}$	$-0.15^{-0.04}_{-0.44}$	$+0.28^{+0.41}_{+0.17}$
48	04 54 11.22	-03 07 59.5	11.1	$3.5^{4.7}_{2.6}$	$-0.06^{+0.06}_{-0.13}$	$+0.07^{+0.20}_{-0.08}$
49	04 54 11.34	-03 01 41.3	11.4	$6.0^{7.4}_{4.3}$	$-0.07^{-0.04}_{-0.14}$	$+0.10^{+0.19}_{+0.01}$
50	04 54 12.38	-03 02 07.0	10.9	$1.6^{1.8}_{1.1}$	$-0.36^{-0.23}_{-0.65}$	$+0.57^{+0.75}_{+0.32}$
51	04 54 12.46	-03 02 17.4	10.9	$2.3^{2.8}_{2.0}$	$-0.09^{+0.02}_{-0.29}$	$+0.08^{+0.30}_{-0.15}$
52	04 54 12.50	-03 01 08.2	11.3	$4.3^{6.0}_{3.4}$	$-0.05^{+0.03}_{-0.11}$	$+0.44^{+0.54}_{+0.35}$
53	04 54 12.75	-02 56 13.5	11.1	$4.2^{5.4}_{3.1}$	$-0.11^{-0.04}_{-0.19}$	$+0.42^{+0.52}_{+0.31}$
54	04 54 13.10	-03 03 11.2	10.7	$1.7^{2.1}_{1.4}$	$+0.07^{+0.16}_{-0.09}$	$+0.43^{+0.63}_{+0.22}$
55	04 54 13.70	-03 02 22.3	11.0	$2.5^{2.9}_{2.3}$	$+0.13^{+0.22}_{+0.01}$	$+0.03^{+0.16}_{-0.10}$
56	04 54 14.01	-03 02 00.0	10.4	$0.9^{1.0}_{0.8}$	$-0.10^{-0.05}_{-0.15}$	$+0.66^{+0.84}_{+0.39}$
57	04 54 14.16	-02 59 29.5	11.5	$6.2^{8.1}_{5.1}$	$-0.19^{-0.10}_{-0.34}$	$+0.25^{+0.35}_{+0.16}$
58	04 54 14.29	-03 00 41.0	11.2	$4.7^{6.5}_{3.5}$	$-0.04^{+0.05}_{-0.11}$	$+0.10^{+0.21}_{-0.02}$
59	04 54 14.34	-03 01 32.4	11.0	$3.5^{4.5}_{3.0}$	$+0.15^{+0.25}_{+0.08}$	$+0.41^{+0.52}_{+0.30}$

Continued on next page

Table 5.3 – *Continued from previous page*

No.	RA	DEC	$\log(M_*/M_\odot)$	Age (Gyr)	[Fe/H]	[Mg/Fe]
60	04 54 14.50	-03 05 13.3	10.1	$8.0^{9.2}_{6.3}$	$-0.50^{-0.41}_{-0.53}$	$+0.98^{+1.02}_{+0.91}$
61	04 54 14.62	-03 01 39.2	10.9	$4.5^{5.5}_{3.5}$	$-0.10^{+0.01}_{-0.15}$	$+0.31^{+0.40}_{+0.22}$
62	04 54 15.16	-03 00 31.4	10.7	$4.2^{5.6}_{3.2}$	$-0.12^{+0.03}_{-0.22}$	$+0.10^{+0.25}_{-0.06}$
63	04 54 15.23	-03 01 07.6	10.6	$3.2^{4.6}_{2.3}$	$+0.03^{+0.17}_{-0.04}$	$+0.15^{+0.28}_{+0.01}$
64	04 54 15.30	-03 01 57.5	11.5	$4.4^{5.3}_{3.7}$	$+0.16^{+0.23}_{+0.08}$	$+0.03^{+0.12}_{-0.06}$
65	04 54 15.95	-03 06 05.0	11.4	$3.0^{3.5}_{2.8}$	$+0.20^{+0.29}_{+0.06}$	$+0.05^{+0.15}_{-0.05}$
66	04 54 16.01	-02 58 56.4	10.4	$1.9^{2.3}_{1.7}$	$-0.06^{+0.01}_{-0.18}$	$+0.25^{+0.46}_{+0.02}$
67	04 54 16.14	-03 00 40.8	10.9	$3.8^{5.0}_{2.9}$	$-0.05^{+0.03}_{-0.11}$	$+0.31^{+0.41}_{+0.21}$
68	04 54 16.31	-03 06 58.6	10.7	$2.9^{4.3}_{2.3}$	$-0.11^{+0.06}_{-0.29}$	$-0.02^{+0.13}_{-0.18}$
69	04 54 16.80	-03 01 31.2	11.0	$1.0^{1.1}_{0.9}$	$-0.13^{-0.07}_{-0.18}$	$+0.41^{+0.65}_{+0.13}$
70	04 54 17.40	-03 02 15.3	10.9	$2.5^{3.3}_{2.1}$	$-0.08^{+0.05}_{-0.33}$	$+0.14^{+0.30}_{-0.04}$
71	04 54 17.73	-02 59 23.0	10.7	$1.0^{1.0}_{0.8}$	$-0.18^{-0.15}_{-0.21}$	$+0.17^{+0.37}_{-0.04}$
72	04 54 17.85	-03 04 21.0	10.0	$5.0^{6.8}_{3.3}$	$-0.12^{+0.00}_{-0.24}$	$+0.45^{+0.61}_{+0.29}$
73	04 54 17.88	-03 01 47.1	11.6	$5.7^{7.5}_{4.8}$	$-0.00^{+0.06}_{-0.05}$	$+0.31^{+0.39}_{+0.23}$
74	04 54 18.65	-03 01 03.3	10.9	$1.2^{6.2}_{1.1}$	$-0.08^{-0.05}_{-0.12}$	$+0.54^{+0.73}_{+0.34}$
75	04 54 18.74	-03 02 37.0	10.8	$4.0^{5.1}_{3.1}$	$-0.11^{-0.05}_{-0.18}$	$+0.36^{+0.46}_{+0.26}$
76	04 54 18.76	-03 00 55.3	11.5	$1.9^{2.0}_{1.7}$	$-0.11^{-0.07}_{-0.15}$	$+0.28^{+0.39}_{+0.17}$
77	04 54 19.01	-03 08 25.7	10.6	$3.5^{4.7}_{2.2}$	$-0.48^{-0.30}_{-0.60}$	$-0.07^{+0.11}_{-0.20}$
78	04 54 19.18	-02 59 25.4	10.8	$1.1^{1.2}_{0.9}$	$-0.13^{-0.08}_{-0.18}$	$+0.28^{+0.54}_{-0.01}$
79	04 54 19.99	-03 01 35.9	10.0	$1.3^{1.8}_{1.1}$	$-0.06^{+0.07}_{-0.46}$	$-0.33^{-0.04}_{-0.55}$
80	04 54 20.40	-03 02 29.2	11.0	$4.2^{8.1}_{2.7}$	$-0.19^{+0.01}_{-0.37}$	$+0.42^{+0.56}_{+0.29}$
81	04 54 22.11	-03 02 49.6	11.1	$2.4^{3.0}_{2.2}$	$+0.17^{+0.24}_{+0.03}$	$-0.04^{+0.07}_{-0.17}$
82	04 54 22.31	-03 02 32.1	10.6	$7.3^{10.4}_{5.8}$	$-0.33^{-0.21}_{-0.46}$	$+0.27^{+0.37}_{+0.17}$
83	04 54 23.13	-03 05 00.7	10.9	$2.6^{2.9}_{2.3}$	$+0.00^{+0.12}_{-0.09}$	$+0.14^{+0.27}_{-0.00}$
84	04 54 24.20	-03 03 38.6	11.4	$3.3^{4.3}_{2.9}$	$+0.10^{+0.20}_{+0.04}$	$+0.27^{+0.35}_{+0.19}$
85	04 54 24.44	-03 03 51.0	11.1	$3.1^{3.4}_{2.8}$	$+0.17^{+0.29}_{+0.05}$	$+0.12^{+0.22}_{+0.01}$
86	04 54 34.07	-03 03 20.7	10.9	$3.8^{7.8}_{1.9}$	$-0.15^{+0.06}_{-0.37}$	$+0.33^{+0.50}_{+0.15}$
87	04 54 36.20	-03 07 03.8	8.9	$3.2^{4.1}_{2.3}$	$-0.53^{-0.20}_{-0.70}$	$+0.80^{+1.07}_{+0.50}$
88	04 54 39.55	-03 05 03.4	11.3	$1.9^{2.2}_{1.8}$	$+0.19^{+0.25}_{+0.07}$	$-0.06^{+0.07}_{-0.19}$
89	04 54 41.44	-03 07 50.0	11.1	$3.9^{5.2}_{2.8}$	$-0.09^{-0.02}_{-0.18}$	$+0.32^{+0.43}_{+0.21}$
90	04 54 45.11	-03 06 25.8	9.6	$2.3^{2.8}_{1.7}$	$-0.56^{-0.30}_{-0.79}$	$+0.49^{+0.80}_{+0.18}$
91	04 54 45.74	-03 02 02.9	11.6	$5.0^{6.8}_{4.0}$	$-0.01^{+0.07}_{-0.06}$	$+0.28^{+0.39}_{+0.18}$
92	04 54 47.32	-03 06 20.1	11.0	$3.4^{4.6}_{2.6}$	$-0.12^{+0.01}_{-0.22}$	$+0.24^{+0.36}_{+0.10}$

APPENDIX

5.A Choosing the best combination of response functions

Although the FSPS code is able to measure many elemental abundances simultaneously, it requires more computation time to measure more elements. Every additional element requires one additional spectral response function. To achieve reliable measurements of [Mg/Fe] within a reasonable computation time, we tested several combinations of response functions to be included in the fitting models. We performed these tests on two sets of data. First, we used our fitting code to fit the high-S/N, stacked spectra of Choi, Conroy, et al. (2014). We then compared our results to the Choi, Conroy, et al.'s measurements. Second, we used the full-spectrum fitting code `alf` (Conroy, Villaume, et al., 2018) to re-measure abundances for a subset of our sample. We compared the `alf` measurements to the results described according to the method described in Section 5.3.

Comparison with Choi14

In this section, we test our measurements of age and metal abundance on 10 high-S/N stacked spectra from Choi, Conroy, et al. (2014). The spectra were stacked from individual spectra in the AGN and Galaxy Evolution Survey (AGES, Kochanek et al., 2012) based on their redshifts, $z = 0.3$ to 0.7, and masses, $M_* = 10^{10.2}$ to $10^{11.3} M_\odot$. The spectral resolution is 6 Å with the wavelength range of 4000 to 5500 Å.

We consider 5 spectral models with different sets of response functions in addition to the four baseline parameters: age, [Z,H], velocity dispersion, and redshift. The combinations are

1. Mg only,
2. alpha elements (Mg, Si, Ca, Ti, and O) all fixed to the same value,
3. Mg and O,
4. Mg and N,
5. "full" combination having the same list of 9 elements as Choi, Conroy, et al. (2014): Mg, N, Fe, O, C, N, Si, Ca and Ti.

We compare the results from each model with the measurements from Choi, Conroy, et al. (2014) (see Figure 5.6). We find that three of five the models agree within 0.1 dex with the measurements from Choi, Conroy, et al. These models are those that include the response functions of (3) Mg and O, (4) Mg and N, and (5) the "full" combination.

In terms of the age measurements, all models perform equally well. The agreement holds when the age is older than $\sim$ 3.5 Gyr. For younger populations, the age measurements obtained here are $\sim$ 0.05 dex older than the ages reported by Choi, Conroy, et al. This trend of the differences in the age measurements is the same as what we found in Chapter 4 when we fit for only ages and [Z/H]. We attributed the difference to the lack of age-sensitive higher-order Balmer lines (<4000Å) in the spectra and possibly the differences in the models used (see Appendix of Chapter 4).

However, we find agreement in [Fe/H] and [Mg/Fe] with Choi, Conroy, et al. only in the three combinations of response functions mentioned above. When the enhancements of all alpha elements are fixed to the same value or when only the [Mg/Fe] response function is included, the [Fe/H] is generally slightly over-predicted while the [Mg/H] is under-predicted (red triangles or orange squares in Figure 5.6). The former case (all alpha elements fixed to the same value) is not surprising because alpha elements generally do not have the same values, and they do not always track each other. Conroy, Graves, et al. (2014) found that Mg tracks O closely, but Mg does not track heavier α elements such as Si, Ca, and Ti. The latter case (Mg response function only) is more interesting. We cannot measure [Mg/Fe] or [Fe/H] well when we only include the response function of Mg. This is probably because the Mg enhancement not only strengthens the absorption lines in the Mg b λ5170 region but also weakens the absorption lines at $\sim$ 4000–4400 Å. The wavelength range also responds to the enhancement of Fe, O, and N (See Figures 7 and 8 of Conroy, Villaume, et al. 2018). For this reason, the agreements improve to within 0.1 dex when we include the response functions of either O or N.

In summary, we found that our fitting code yields results that agree reasonably well (within 0.1 dex) with the literature when the we use the following models: Mg and N, Mg and O, or the "full" combination. However, the "full" combination model, which includes 13 parameters, is costly in terms of the computational time. Therefore, it is difficult to apply to a large sample. Thus, we restrict further consideration to the two simpler models that include the response functions of either Mg and N or Mg

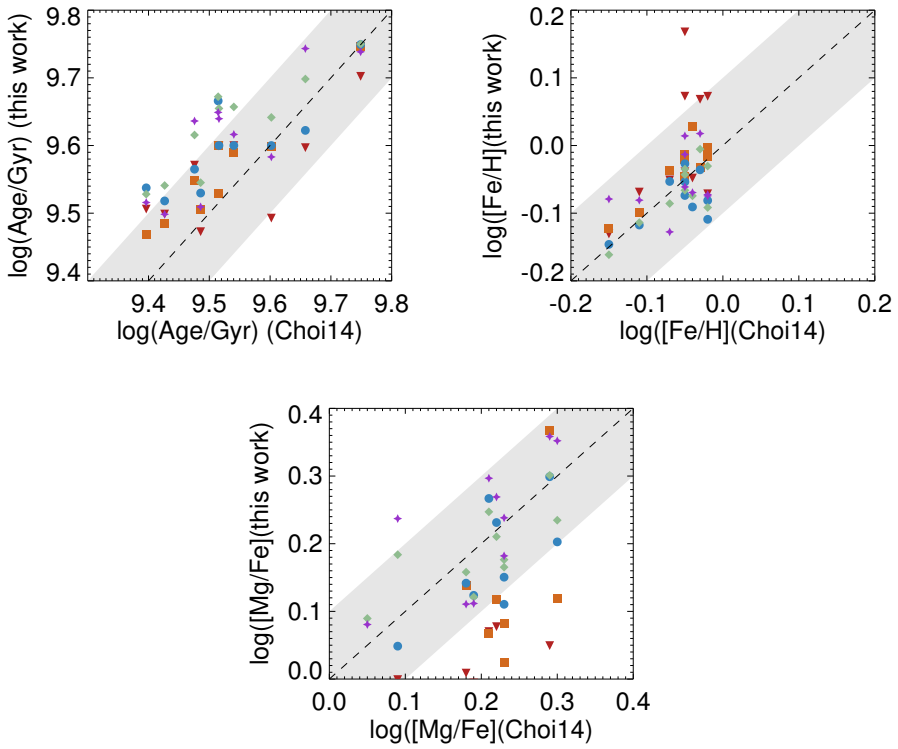

Figure 5.6: Comparison with the measurements of Choi, Conroy, et al. (2014) using different combinations of metal elements in the response functions. The combinations are (1) [Mg/Fe] only (orange squares), (2) alpha elements [Mg, Si, Ca,Ti , O/Fe] all fixed to the same value (red downward triangles), (3) [Mg/Fe] and [O/Fe] (blue circles), (4) [Mg/Fe] and [N/Fe] (green diamonds), and (5) Mg, N, Fe, O, C, N, Si, Ca and Ti/Fe (purple stars). The latter three combinations show the best agreement with Choi, Conroy, et al. (2014).

and O.

Comparison to the `alf` full-spectrum fitting code

In order to choose the better of the remaining two models, we compare the results from our fitting code to the results from the publicly available "absorption line fitting code" (`alf`, Conroy, Villaume, et al., 2018). `alf` is a fitting algorithm that uses the Markov Chain Monte Carlo method of parameter estimation. It constructs empirical SSP spectra from the MIST isochrones (Choi, Dotter, et al., 2016), the MILES and Extended IRTF spectra libraries, and theoretical response functions of individual elements based on the Kurucz suite of routines. Our code uses a different set of empirical SSP spectra (Conroy, Gunn, and M. White, 2009) but uses the same response functions. All of the `alf` results presented here were modeled with the

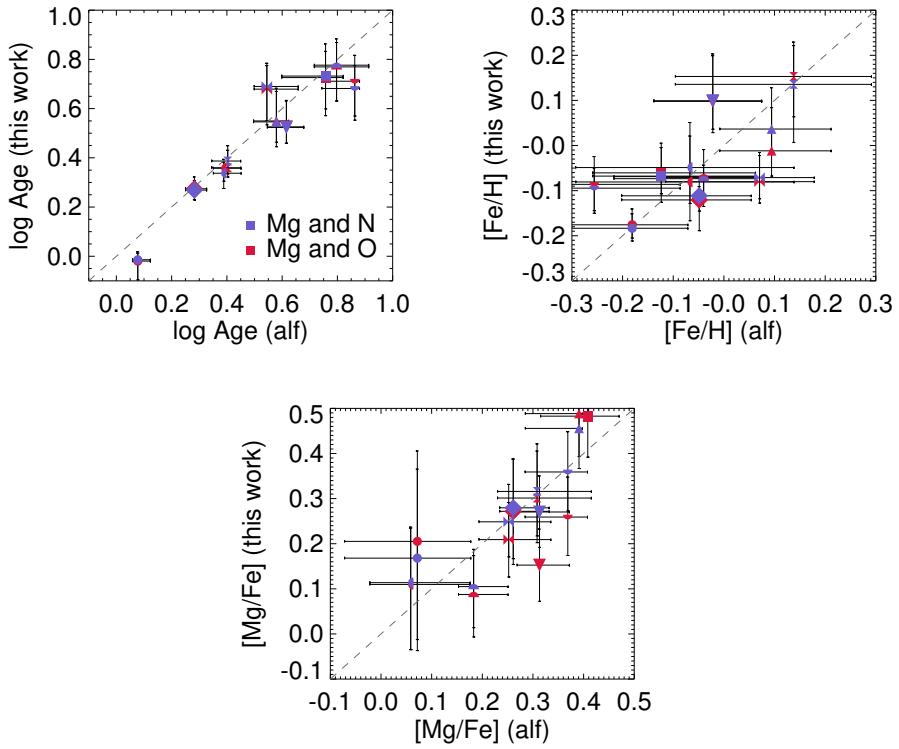

Figure 5.7: Comparison of parameters of 10 galaxies measured with `alf` (Conroy, Villaume, et al., 2018), which fits for the abundances of nine individual elements, and measure with our code using the combination of response functions of (1) [Mg/Fe] and [N/Fe] and (2) [Mg/Fe] and [O/Fe]. We chose combination (1) in our final fitting code to apply to the full sample.

Kroupa IMF and the "simple" mode of fitting, which fits for 13 parameters including stellar age, [Z/H], and abundances of Fe, C, N, O, Na, Mg, Si, Ca, and Ti.

Because the `alf` computational time for a single spectrum is long ($\gtrsim$ 100 cpu hours), we selected the 10 highest S/N spectra from our sample in MS0451 ($z \sim 0.54$). Figure 5.7 shows the comparison between the two methods. Our measurements are remarkably consistent with the results from `alf`, given that our code is simpler. However, we found that in some of the cases, the model that includes the response functions of Mg and N yields better agreement, especially in terms of [Mg/Fe], than the model that includes the response functions of Mg and O. Based on the results in this section, we chose the model that includes the response functions of Mg and N to apply to our data.

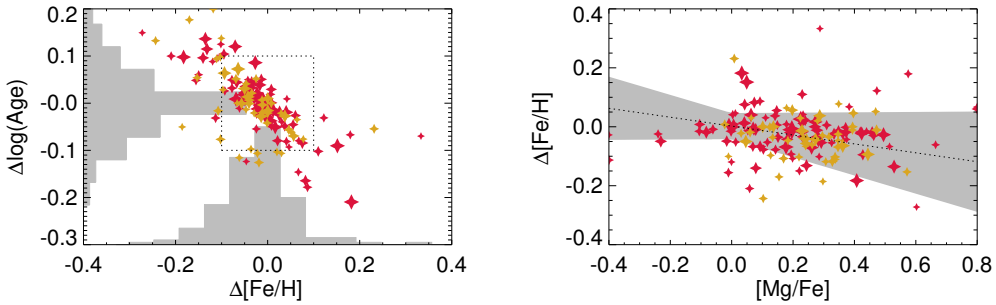

Figure 5.8: Comparison of the age and [Fe/H] measurements using the models in Chapter 4 (same models but without α enhancement) and the models used here. The uncertainties in both parameters lie along the age-metallicity degeneracy direction with a slight dependence on [Mg/Fe].

Comparison between the fitting method in Chapter 4 and in this work

Here we compare the [Fe/H] and age measurements derived from the method in Chapter 4 to those derived in this paper. In Chapter 4, we masked out the Mg b absorption region in the spectra and measured [Fe/H] and age with without α enhancement. In this paper, we fit the full spectra with models that treat N and Mg abundances as free parameters. We use both methods to fit the higher-redshift spectra in this paper.

Figure 5.8 shows the comparisons of the results from both methods. The x- and y-axes represent the results from this paper subtracted from the results from Chapter 4. The left panel shows that the differences in age and [Fe/H] lies along the age–metallicity degeneracy. The mode of the distribution is zero. I.e., there are no systematic differences in the [Fe/H] nor age measured by either method. Most of the discrepancies are within 0.1 dex in either parameter, as shown by the dotted box. However, the distribution of the differences in measured [Fe/H] (Δ[Fe/H]) is slightly asymmetric, with a heavier distribution toward negative values. This means that we are more likely to have underestimated [Fe/H] in this paper compared to Chapter 4.

The right panel shows Δ[Fe/H] vs. [Mg/Fe] along with the best-fit line. Although the best-fit slope is statistically inconsistent with zero, its best-fit value suggests that when [Mg/Fe] is high, we possibly underestimate [Fe/H]. The reason for this is also likely due to the response function of Fe and Mg. The spectrum in the 4000–4400 Å range responds to Mg and Fe in the opposite direction. Therefore, the models without elemental enhancements are likely to underestimate the [Fe/H] when fit to a Mg-enhanced spectrum.

5.B Chemical evolution model with enriched inflow

If the infalling gas has a metallicity of $Z_{\rm inf}$, we can modify equation 5.4 to

$$dM_{Z,g} = ydM_* - Z_g dM_* - \frac{\eta}{1-R} Z_g dM_* + Z_{\rm inf} dM_{\rm inf} \tag{5.8}$$

where $dM_{\rm inf}$ is the mass inflow rate. We can substitute $Z_g dM_*$ with $dM_{Z,*}$, integrate the equation, and divide all terms by dM_* to get

$$Z_g r_g = y - (1 + \frac{\eta}{1-R}) Z_* + < Z_{\rm inf} > \frac{M_{\rm inf}}{M_*} \tag{5.9}$$

Assuming the galaxy started with no gas, we can estimate the total infall mass for quiescent galaxies as the current stellar mass plus the total outflow mass: $M_{\rm inf} = (1 + \frac{\eta}{1-R}) M_*$. The equation for quiescent galaxies is then

$$Z_* \approx \frac{y}{1 + \frac{<\eta>}{1-R}} + < Z_{\rm inf} > \quad \text{or} \quad < \eta > \approx \left(\frac{y}{Z_* - < Z_{\rm inf} >} - 1 \right) (1 - R) \tag{5.10}$$

The equation above suggests that we would underestimate the mass-loading factor by assuming that the inflowing gas in pristine. The amount of the underestimation is shown in Figure 5.9. The y-axis is the ratio between the η derived assuming pristine inflow to the η derived from the model with enriched inflow. The amount of the underestimation depends on both inflow and stellar metallicity. For simplicity in quantifying the amount of the underestimation, we parameterize the inflow metallicity as a fixed fraction of the final stellar metallicity: $Z_{\rm inf} = \alpha Z_*$. (This is similar to the parameterization used by Finlator et al. 2008 but with the stellar metallicity instead of the gas-phase metallicity.)

The underestimation of the mass-loading factor when gas inflow is assumed to be pristine is estimated to be at most 20% on a linear scale. Studies of infalling H I absorbing complexes generally have found that H I filaments and the circum-galactic medium are metal-poor, $\log Z/Z_\odot < -1$ (e.g., Churchill et al., 2012; Hafen et al., 2017). Thus, based on Figure 5.9, even in a very enriched inflow case at the high-mass end ($Z_{\rm inf} = 0.1 Z_* \sim -1$ dex), the estimated η would be at most incorrect by 20%, or less than 0.1 dex.

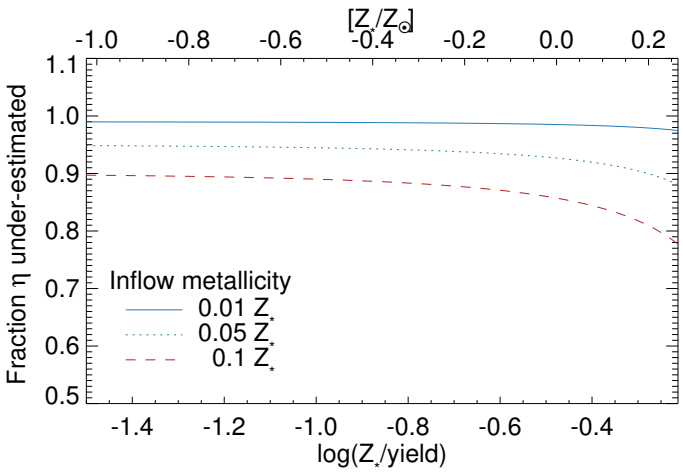

Figure 5.9: The underestimation of the mass-loading factor (the ratio between η derived assuming pristine infalling gas to η derived with enriched inflow) as a function of stellar metallicity. The bottom x-axis is the metallicity in units of the yield. The top x-axis shows the metallicity when the yield is assumed to be $3Z_{\odot}$. Each line corresponds to a different assumed metallicity of the inflow.

Chapter 6

CONCLUSION AND FUTURE WORK

In this thesis, we use spectroscopic observations to investigate galaxies at different epochs from the age of reionization to the local universe. The sample in this thesis includes both star-forming and quiescent galaxies. We focus on the effect of galactic feedback on galaxy evolution. The summary of this thesis is as follows.

6.1 First epoch: Escape fractions in the early universe is likely large enough for galaxies to be the primary agents in reionizing the universe.

In Chapter 2, we first explored galaxies in the early universe, $z \sim 5$, to understand the role of star-forming galaxies in reionizing the universe. Specifically, we attempted to quantify the fraction of the ionizing photons that escape from star-forming galaxies into the intergalactic medium, a.k.a. the escape fraction ($f_{\rm esc}$). The approach was to use absorption line spectroscopy to estimate the $f_{\rm esc}$ of four star-forming, gravitationally-lensed galaxies at $z \simeq 5$. The work more than doubled the sample size of previous $f_{\rm esc}$ estimations in T. A. Jones, R. S. Ellis, et al. (2013) and pushed the redshift range to higher than what had been done before.

We found that the escape fractions vary from galaxy to galaxy and appear to have no significant evolution over time. After accounting for reddening and other related assumptions, we inferred a median absolute escape fraction of ionizing UV photons to be greater than $f_{\rm esc,abs} \sim 11\%$. This supports the proposition that galaxies were the primary agents for cosmic reionization. Moreover, we used high gravitational magnification, for the first time, to *spatially resolve* and detect variations in escape fractions within a galaxy. The variation of escape fractions within the galaxy is comparable to that across the integrated analysis of the full sample, which led me to conclude that the escape fraction is governed by small-scale structure.

In order to explain the origin of the wide variation in the escape fractions, we found an anti-correlation between the star-formation rate and the inferred escape fraction. We concluded that the anti-correlation is associated with the delay time after an episode of star formation. Feedback seems to be effective in governing both a low SFR and a smaller HI covering fraction. The delay time allows feedback to clear the covering clouds and is consistent with 'burst-like' star formation in galaxies at

high redshift seen in simulations.

6.2 Second epoch: Galaxies at $z \sim 2$ are more dynamically turbulent and have flatter metallicity gradients than previously thought.

In Chapter 3, we used spatially resolved spectroscopic observations of galaxies at the peak era of star formation to study primitive disks. In particular, we measured their gas-phase metallicity gradients and kinematic properties to provide constraints on the strength of feedback in high-redshift galaxies. We observed 11 gravitationally-lensed star-forming galaxies at $z \sim 2$ with an AO-corrected integral-field spectrograph. The typical spatial resolution in my study is sub-kpc, smaller than a half-light radius of a typical L^* galaxy at $z \simeq 2$. Combining with the sample in Jones et al. 2013, we approximately tripled the existing metallicity gradient measurements in the redshift range.

First, we found a lower fraction of rotationally-supported systems than reported from larger kinematic surveys whose spatial resolutions were coarser. We demonstrated that the discrepancy might be partly due to the improved spatial resolution in my study. We performed a simple experiment by smoothing two kinematically-disturbed galaxies in our sample with seeing typical of a non-AO study. After reanalyzing the kinematics, we found that the two galaxies would be categorized as smooth rotating disk galaxies. Therefore, we concluded that high spatial resolution is crucial for an accurate diagnosis of the kinematic properties and dynamical maturity of $z \simeq 2$ galaxies.

Second, we found a large diversity of metallicity gradients among rotating galaxies at $z \simeq 2$. A much higher fraction of the galaxies has weak metallicity gradients than previously found (T. A. Jones, R. S. Ellis, et al., 2013). The degree of ordered rotation has no obvious effect on the gradient slope. This indicates that the formation and destruction of gradients are on short timescales, comparable to the galaxy dynamical time. Finally, we found a weak correlation between galaxy-integrated metallicity and gradient slope. We interpreted the correlation as an effect of feedback; the same feedback, which reduces the gas-phase metallicity, evidently also flattens the gradients. The diversity of metallicity gradients at $z \simeq 2$ is therefore likely caused in part by differences in the recent feedback history.

6.3 Third epoch: Stellar metallicities of $z < 1$ quiescent galaxies suggest that stellar feedback-driven outflow is a power-law function of galaxy mass.

In the second half of my thesis, we focus on the stellar metallicity of quiescent galaxies at $z < 1$. Because stellar metallicities reflect the "star formation history-averaged" metallicity from the whole galaxy, it is less susceptible than the gas to instantaneous fluctuations. Obtaining stellar metallicities over a range of galaxy mass (stellar mass–metallicity relation, MZR) can directly reveal the overall feedback on galactic chemical evolution.

In Chapter 4, we measured metallicities of 62 quiescent galaxies in a galaxy cluster at $z \sim 0.4$, in terms of iron abundances ([Fe/H]), and compared to those of local SDSS quiescent galaxies. The metallicities were measured using the full-spectrum fitting technique (Conroy, Gunn, and M. White, 2009). We detected for the first time an evolution of the MZR with redshift. Moreover, we found that the slopes of the MZRs at both redshifts are generally shallow. The MZR slope indicates the dependence of feedback on galaxy mass and redshift. If lower-mass galaxies have relatively stronger feedback, they would lose higher fractions of produced metals, and the slope of the MZR would be steep. The observed shallow slopes suggest that the relative galactic feedback strength, as indicated by the mass-loading factor (defined as outflow divided by SFR), might be weakly dependent on mass over the observed mass range.

We also estimated the stellar MZRs of quiescent galaxies at the time of their formation. This technique traces quiescent galaxies observed at different redshifts back to their formation epoch by using measurements of galaxy age. We showed that the evolution of the stellar MZR with observed redshift is easily explained by an underlying evolution with their formation time. The technique can potentially be used to (1) bypass the biases in measuring stellar metallicities in young galaxies (Ge et al. 2018); (2) to compare the stellar MZRs of quiescent galaxies with the gas-phase MZRs of star-forming galaxies at different redshifts; and (3) to determine whether quiescent galaxies evolve passively after quenching.

In Chapter 5, we expand the sample to include 92 more galaxies in a galaxy cluster at $z \sim 0.55$. We additionally measure magnesium (Mg) enhancement via the inclusion of the Mg response function (Conroy, Villaume, et al., 2018) in the spectral models. Mg is mainly produced in core-collapse supernovae. It has a shorter recycling timescale than Fe, which is predominantly produced by Type Ia supernovae. Surprisingly, when we plot [Mg/H] (instead of [Fe/H]) as a function of

galaxy mass, the evolution with redshift disappears. The evolution of [Fe/H] seen in Chapter 4 is likely a result of the longer recycling timescale for Fe compared to Mg. The Fe abundance appeared to be evolving because older galaxies have shorter formation timescales.

By using Mg as an indicator of metals, the recycling timescale can be approximated as instantaneous. We were able to use galactic chemical evolution model to archaeologically quantify the mass-loading factor as a function of mass from the mass–Mg relation. The measured mass-loading factors are consistent with the results of direct observational methods for outflow measurements. We find that the average mass-loading factor is a power-law function of galaxy stellar mass, $\eta \propto M_*^{-0.21\pm0.09}$. The measured scaling relation is consistent with an analytic prediction in which outflow is caused by star-formation feedback in a turbulent disk (Hayward et al., 2017).

In summary, we found that star-formation feedback is effective in creating porosity in galaxies in the early universe, high enough for galaxies to be the primary agents in ionizing the universe. Feedback is also likely responsible for the flat metallicity gradients and the turbulent kinematics in star-forming galaxies at $z \sim 2$. Lastly, we determine the dependence of the strength of galactic feedback on galaxy mass and redshift via the construction of the stellar MZR. We found that the mass loading factor is a function of galaxy mass over the observed mass range of $10^{9.7} < M_*/M_\odot < 10^{11.5}$.

6.4 Future Work

To understand how the first galaxies formed and how they subsequently evolved into systems in the local universe, it is important to trace the origin of matter and chemical elements back to the earliest times in the universe. The upcoming launch of the James Webb Space Telescope (JWST) will enable expansion of the work in this thesis into higher redshifts. For example, the measurements of escape fractions of galaxies in the epoch of reionization ($z > 6$) will be possible with the JWST/NIRSpec spectroscopy in the near-infrared. The JWST/NIRISS slitless spectrograph, with its sensitivity and spatial resolution, can be used to observe large numbers of rest-frame optical spectra of galaxies up to $z \sim 4$. These observations will enable the measurements of both metallicity gradients (in gravitationally-lensed star-forming arcs) and stellar metallicities. We will be able to trace the build-up of stellar mass, metals, interstellar dust, and massive black holes at $z \sim 4$ and beyond for the first time.

In the meantime, there are many projects that can be done and are needed to be done to prepare for those observations. One particular example is to improve the techniques used to measure metallicities in galaxies. For stellar metallicities, it is currently known that measuring stellar metallicities of young galaxies (< 1 Gyr old) is prone to large biases (e.g. Leethochawalit, Kirby, et al., 2018; Ge et al., 2018; Cid Fernandes, 2018). This is because young stars have weak or no absorption lines. The age of the universe at redshift $4 - 2$ is only $1.5 - 3$ Gyr. Therefore, the measurements of stellar metallcities of any galaxies at this redshift range will run into this problem. A possible solution is to model the aggregated spectra of stars and gas together, from the rest-frame UV to optical wavelengths (e.g. Steidel, Strom, et al., 2016).

Another aspect is to check the possible discrepancies between the stellar metallcities measured from resolved stellar population and those from integrated spectra. We present an ongoing work using Palomar/CWI observations here.

Resolving the Discrepancy in the Stellar Mass–Metallicity Relation

Context: Galactic metals are measurable in two forms: gas phase and stellar phase. Gas-phase metallicity reflects the current concentration of metals available for present star formation. Stellar metallicity indicates the amount of metals incorporated into stars at the time they formed. Stellar metallicity therefore reflects the 'star-formation history average' metal abundance and is less susceptible to instantaneous fluctuations. In a recent paper (Leethochawalit et al. 2019), we demonstrated that we can use stellar metallicity to archaeologically derive an average outflow strength, in terms of mass-loading factors (η), that each quiescent galaxy experienced. Our measured η's are consistent with the results from other observational methods for direct outflow measurements, but our approach allows us to measure past outflows for many (> 100) galaxies at $z \sim 0.5$.

Generally, two independent techniques are used to derive stellar metallicities depending on the distance of galaxies. Observable satellite dwarf galaxies ($M_* \lesssim 10^9 M_\odot$) are mainly local and resolved into stars. Their metallicities are therefore derived from individual stars, either spectroscopically or using a color-magnitude diagram (CMD). On the other hand, metallicities of more massive galaxies ($M_* \gtrsim 10^9 M_\odot$) are derived from their integrated stellar light since they are too far for individual stars to be resolved. Regardless of technique and mass range, galaxies with higher stellar masses are more metal-rich than galaxies with lower stellar masses. The

correlation, the stellar mass–metallicity relation (MZR), holds true for both dwarf and more massive galaxies.

However, there is increasing observational evidence that the MZR of dwarf galaxies has a steeper slope and lower normalization than expected from extrapolating the MZR of more massive galaxies to galaxies of lower mass. For example, consider the population of quiescent galaxies. Kirby, Cohen, et al. (2013) derived the stellar MZR of Local Group dwarf galaxies with masses from 10^3 to $10^9 M_\odot$, based on the measurements of metallicities of individual stars. The MZR follows a universal single power-law relation

$$[\mathrm{Fe/H}] = (-0.79 \pm 0.06) + (0.30 \pm 0.02) \log\left(\frac{M_*}{10^9 M_\odot}\right). \tag{6.1}$$

All quiescent and star-forming Local Group dwarf galaxies follow this relation.

On the other hand, several works consistently found that the MZR of more massive $z \sim 0$ quiescent galaxies has a shallower slope. Gallazzi, Bell, et al. (2014), Choi, Conroy, et al. (2014), Conroy, Graves, et al. (2014), and Leethochawalit, Kirby, et al. (2018) used different fitting codes and different integrated stellar spectral models to measure stellar metallicities from integrated-light spectra of $M_* \gtrsim 10^9 M_\odot$ quiescent galaxies. All found that the MZR in this mass range follows a power-law relation but with a slope of < 0.15 dex per log mass. The normalization is -0.20 ± 0.03 at $10^9 M_\odot$. The slope is roughly half, and the normalization is ~ 0.6 dex higher, compared to the same quantities derived from individual stars in Local Group dwarf galaxies (see Figure 6.1).

One possible resolution for the apparent discrepancy in the MZRs of the two mass ranges is that there is a transition in the MZR around $10^8 - 10^9 M_\odot$. The existing measurements in this mass range are scarce. Only one Local Group dwarf galaxy in the sample of Kirby, Cohen, et al. (2013) is more massive than $10^8 M_\odot$, while none of the studies based on integrated light spectra have galaxies with stellar mass below $10^9 M_\odot$ in their sample. If the MZR does transition at $10^{8-9}\ M_\odot$, this would suggest that galaxies of the two mass ranges have different feedback mechanisms. Theoretically, it has been suggested that energy-driven winds ($\eta \propto v_c^{-2}$) dominate in low-mass galaxies (Murray et al., 2005), while momentum-driven winds ($\eta \propto v_c^{-1}$) dominate in high mass galaxies (Hopkins, Quataert, et al., 2012).

Systematics in integrated-light vs. resolved stellar spectroscopy: However, there is a more concerning explanation for the discrepancy in the MZR: a systematic difference between the metallicities measured from individual stars and those measured

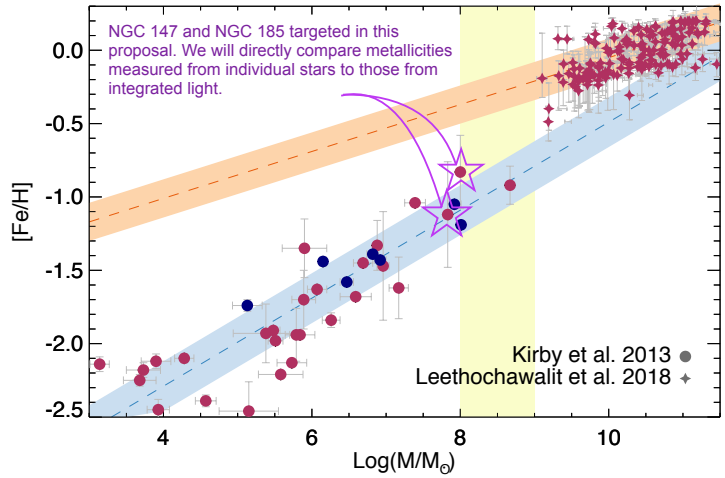

Figure 6.1: Current MZR of $z \sim 0$ galaxies. The measurements below $10^8 M_\odot$ are from resolved stellar spectroscopy of Local Group dwarfs (solid circles, Kirby, Cohen, et al., 2013), while the measurements of more massive galaxies (diamonds) are from Leethochawalit, Kirby, et al. (2018). Blue and red colors represent star-forming and quiescent galaxies, respectively. The extrapolations from the best-fit lines of the two populations to the entire range of galaxy mass (dashed lines) are incompatible. We first plan to directly compare ages and metallicities measured from integrated light spectra of NGC 147 and NGC 185 dwarf galaxies (purple stars) to those from resolved stellar spectroscopy.

from integrated light. Directly comparing integrated-light and stellar abundances has been observationally challenging because there are subtleties and complications associated with "integrating" the light of galaxies that are close enough for resolved stellar spectroscopy. Multiple works have tested the consistency between star formation history (SFH), metallicity, and age derived from integrated light and those from other methods. Although these tests yield reasonable consistency, their samples are limited to simple stellar populations such as star clusters, and the comparison measurements are mainly limited to results from CMDs (e.g. González Delgado et al., 2010; Barber et al., 2014; Conroy, Graves, et al., 2014).

Recent work by Ruiz-Lara et al. (2018) suggests that this concern deserves attention, especially in more complex systems such as galaxies. Ruiz-Lara et al. obtained an integrated spectrum of Leo A, one of the the Local Group's star-forming dwarf irregular galaxies with $M_* \sim 3 \times 10^6 M_\odot$. They scanned a slit spectrograph on the 10m GTC telescope through 16 sky positions. The total integration time was 16 hrs (1 hour at each position) and the final S/N of the combined spectrum is ~ 100 Å^{-1}. They derived the SFH and the metallicity of each SFH age bin from the

integrated spectrum. *They found that while the SFH is consistent with that derived from the CMD, the derived [Fe/H] at all age bins are higher than the average value of individual stars obtained by Kirby, Rizzi, et al. (2017).* The differences in [Fe/H] range from ~ 0.25 dex in the intermediate age bins to > 0.6 dex in younger age bins.

If the discrepancy in the measured [Fe/H] of younger stellar population is truly as large as 0.6 dex, it potentially casts doubt on all existing MZRs measured in higher-mass galaxies based on integrated light. This is mainly because the light from younger stellar populations (and hence more massive stars) tends to dominate the integrated light of galaxies. Due to the S/N of their data, these works either assume that each galaxy is a single stellar population (SSP, e.g., Choi, Conroy, et al., 2014; Leethochawalit, Kirby, et al., 2018; Sybilska, Kuntschner, et al., 2018) or assume a fixed [Fe/H] at all age bins of SFH (Gallazzi, Bell, et al., 2014). This means that their results are closer to light-weighted [Fe/H], which is more sensitive to younger population and have greater discrepancy with the individual star measurements.

However, there can be other less alarming reasons that explain the discrepancy found by Ruiz-Lara et al. First, Leo A is still star-forming and therefore emits nebular emission lines. If their modeling of the Balmer emission lines from the interstellar medium was not perfect, it could have affected the depths of the Balmer absorption lines, which measure the stellar population age. In turn, this could have affected the measured metallicity according to the age–metallicity degeneracy. Old, quiescent galaxies would not suffer the same source of uncertainty. Second, Leo A is metal poor ([Fe/H] ~ -1.5). Most stellar libraries that are used in the modelling of integrated light spectra do not contain many stars with [Fe/H] < -1. If the incompleteness of the metal-poor stars in the stellar libraries causes the discrepancy found by Ruiz-Lara et al., it will not affect existing MZRs of higher mass galaxies since those galaxies have higher metallicities.

Method: We plan to take advantage of the large field of Palomar/CWI to further investigate the systematic difference between metallicities measured from individual stars and from integrated light. This can be done by measuring stellar metallicities from both integrated light and individual stars of two dwarf elliptical (dE) galaxies: NGC 147 and NGC 185. Both are satellites of M31, whose [Fe/H] measurements based on resolved stellar spectroscopy already exist (Kirby, Cohen, et al., 2013). To our knowledge, this will be the second paper, aside from Ruiz-Lara et al., to perform such a comparison in galaxies and will be the first for quiescent galaxies.

The preceding proposal for this project was submitted in April 2019 to the Caltech

Optical Observatories for review.

BIBLIOGRAPHY

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. (2003), ApJ 591, 499

Adams, W. S. (1949), ApJ 109, 354

Alexandroff, R. M., Heckman, T. M., Borthakur, S., Overzier, R., & Leitherer, C. (2015), ApJ 810, 104, 104

Anglés-Alcázar, D., Davé, R., Özel, F., & Oppenheimer, B. D. (2014), ApJ 782, 84, 84

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. (2009), Annual Review of Astronomy and Astrophysics 47, 481

Atek, H. et al. (2015), ApJ 814, 69, 69

Auger, M. W., Treu, T., Brewer, B. J., & Marshall, P. J. (2011), MNRAS 411, L6

Barber, C., Courteau, S., Roediger, J. C., & Schiavon, R. P. (2014), MNRAS 440, 2953

Bassett, R. et al. (2019), MNRAS 483, 5223

Becker, G. D. et al. (2015), MNRAS 447, 3402

Becker, R. H. et al. (2001), AJ 122, 2850

Behroozi, P. S., Wechsler, R. H., & Conroy, C. (2013), ApJ 770, 57, 57

Bell, E. F. & de Jong, R. S. (2001), ApJ 550, 212

Belli, S., Jones, T., Ellis, R. S., & Richard, J. (2013), ApJ 772, 141, 141

Bian, F., Kewley, L. J., Dopita, M. A., & Blanc, G. A. (2017), ApJ 834, 51, 51

Bigiel, F., Bolatto, A. D., et al. (2010), ApJ 725, 1159

Bigiel, F., Leroy, A., et al. (2008), AJ 136, 2846

Birnboim, Y. & Dekel, A. (2003), MNRAS 345, 349

Blanton, M. R. & Roweis, S. (2007), AJ 133, 734

Bolatto, A. D. et al. (2013), Nature 499, 450

Borthakur, S., Heckman, T. M., Leitherer, C., & Overzier, R. A. (2014), Science 346, 216

Boselli, A. et al. (2014), A&A 564, A66, A66

Bouché, N. et al. (2012), MNRAS 426, 801

Boutsia, K. et al. (2011), ApJ 736, 41, 41

Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Broadhurst, T. J., & Franx, M. (2004), ApJ 611, L1

Bouwens, R. J., Illingworth, G. D., Franx, M., et al. (2009), ApJ 705, 936

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. (2015), ApJ 811, 140, 140

Bouwens, R. J., Smit, R., et al. (2016), ApJ 831, 176, 176

Bresolin, F. (2011), ApJ 730, 129, 129

Brewer, B. J. et al. (2011), MNRAS 412, 2521

Bromm, V. & Yoshida, N. (2011), Annual Review of Astronomy and Astrophysics 49, 373

Bruzual, G. & Charlot, S. (2003), MNRAS 344, 1000

Bunker, A. J., Moustakas, L. A., & Davis, M. (2000), ApJ 531, 95

Burnham, K. & Anderson, D. (2003), Springer New York

Calvi, R., Poggianti, B. M., Vulcani, B., & Fasano, G. (2013), MNRAS 432, 3141

Calzetti, D. et al. (2000), ApJ 533, 682

Cappellari, M. (2017), MNRAS 466, 798

Castellano, M. et al. (2012), A&A 540, A39, A39

Chabrier, G. (2003), PASP 115, 763

Chisholm, J., Tremonti, C. A., Leitherer, C., & Chen, Y. (2017), MNRAS 469, 4831

Choi, J., Dotter, A., et al. (2016), ApJ 823, 102, 102

Choi, J., Conroy, C., et al. (2014), ApJ 792, 95, 95

Chornock, R. et al. (2014), arXiv e-prints, arXiv:1405.7400, arXiv:1405.7400

Churchill, C. W. et al. (2012), ApJ 760, 68, 68

Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. (2005), MNRAS 358, 363

Cid Fernandes, R. (2018), MNRAS 480, 4480

Conroy, C. & Gunn, J. E. (2010), ApJ 712, 833

Conroy, C. & van Dokkum, P. G. (2012), ApJ 760, 71, 71

Conroy, C., Graves, G. J., & van Dokkum, P. G. (2014), ApJ 780, 33, 33

Conroy, C., Gunn, J. E., & White, M. (2009), ApJ 699, 486

Conroy, C., Villaume, A., van Dokkum, P. G., & Lind, K. (2018), ApJ 854, 139, 139

Cool, R. J. et al. (2008), ApJ 682, 919

Cooper, M. C., Newman, J. A., Davis, M., Finkbeiner, D. P., & Gerke, B. F. (2012), Astrophysics Source Code Library

Cooper, M. C., Tremonti, C. A., Newman, J. A., & Zabludoff, A. I. (2008), MNRAS 390, 245

Crain, R. A. et al. (2015), MNRAS 450, 1937

Cresci, G. et al. (2010), Nature 467, 811

Croton, D. J. et al. (2006), MNRAS 365, 11

Dahle, H. et al. (2013), ApJ 773, 146, 146

Davé, R., Finlator, K., & Oppenheimer, B. D. (2011), MNRAS 416, 1354

Davé, R., Katz, N., Oppenheimer, B. D., Kollmeier, J. A., & Weinberg, D. H. (2013), MNRAS 434, 2645

Davies, F. B. et al. (2018), ApJ 864, 142, 142

Davies, R. L. et al. (2019), ApJ 873, 122, 122

De Lucia, G. & Borgani, S. (2012), MNRAS 426, L61

De Rossi, M. E., Bower, R. G., Font, A. S., Schaye, J., & Theuns, T. (2017), ArXiv e-prints

de Vaucouleurs, G. (1958), ApJ 128, 465

Dekel, A. & Silk, J. (1986), ApJ 303, 39

Dijkstra, M., Haiman, Z., & Loeb, A. (2004), ApJ 613, 646

Dijkstra, M., Gronke, M., & Venkatesan, A. (2016), ApJ 828, 71, 71

Donahue, M. & Shull, J. M. (1987), ApJL 323, L13

Dunlop, J. S. et al. (2013), MNRAS 432, 3520

Dutton, A. A., Conroy, C., et al. (2011), MNRAS 416, 322

Dutton, A. A., van den Bosch, F. C., et al. (2011), MNRAS 410, 1660

Elíasdóttir, Á. et al. (2007), ArXiv e-prints

Erb, D. K., Shapley, A. E., et al. (2006), ApJ 644, 813

Erb, D. K., Quider, A. M., Henry, A. L., & Martin, C. L. (2012), ApJ 759, 26, 26

Erb, D. K., Steidel, C. C., et al. (2014), ApJ 795, 33, 33

Estrada-Carpenter, V. et al. (2019), ApJ 870, 133, 133

Etherington, J. et al. (2017), MNRAS 466, 228

Faber, S. M. (1973), ApJ 179, 731

Faber, S. M. & Jackson, R. E. (1976), ApJ 204, 668

Faber, S. M., Willmer, C. N. A., et al. (2007), ApJ 665, 265

Faber, S. M., Phillips, A. C., et al. (2003), *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*, ed. by M. Iye & A. F. M. Moorwood, vol. 4841, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1657

Faisst, A. L. (2016), ApJ 829, 99, 99

Fan, X., Carilli, C. L., & Keating, B. (2006), Annual Review of Astronomy and Astrophysics 44, 415

Ferrara, A. & Loeb, A. (2013), MNRAS 431, 2826

Filippenko, A. V. & Sargent, W. L. W. (1992), AJ 103, 28

Finkelstein, S. L. et al. (2015), ApJ 810, 71, 71

Finlator, K. & Davé, R. (2008), MNRAS 385, 2181

Fioc, M. & Rocca-Volmerange, B. (1999), ArXiv Astrophysics e-prints

Fitzpatrick, P. J. & Graves, G. J. (2015), MNRAS 447, 1383

Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. (2009), ApJ 706, 1364

Förster Schreiber, N. M., Genzel, R., Lehnert, M. D., et al. (2006), ApJ 645, 1062

Förster Schreiber, N. M., Genzel, R., Newman, S. F., et al. (2014), ApJ 787, 38, 38

Förster Schreiber, N. M., Renzini, A., et al. (2018), ApJS 238, 21, 21

Förster Schreiber, N. M., Shapley, A. E., et al. (2011), ApJ 739, 45, 45

Fumagalli, M. et al. (2016), ApJ 822, 1, 1

Gallazzi, A., Charlot, S., Brinchmann, J., & White, S. D. M. (2006), MNRAS 370, 1106

Gallazzi, A., Bell, E. F., Zibetti, S., Brinchmann, J., & Kelson, D. D. (2014), ApJ 788, 72, 72

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. (2005), MNRAS 362, 41

Garnett, D. R. & Shields, G. A. (1987), ApJ 317, 82

Garnett, D. R. (2002), ApJ 581, 1019

Ge, J. et al. (2018), MNRAS 478, 2633

Geach, J. E. et al. (2014), Nature 516, 68

Genzel, R., Burkert, A., et al. (2008), ApJ 687, 59

Genzel, R., Förster Schreiber, N. M., et al. (2014), ApJ 796, 7, 7

Genzel, R., Newman, S., et al. (2011), ApJ 733, 101, 101

Genzel, R., Tacconi, L. J., et al. (2006), Nature 442, 786

Gerke, B. F. et al. (2007), MNRAS 376, 1425

Giavalisco, M. et al. (2004), ApJL 600, L93

Gibson, B. K., Pilkington, K., Brook, C. B., Stinson, G. S., & Bailin, J. (2013), A&A 554, A47, A47

Gillman, S. et al. (2019), MNRAS 486, 175

Gladders, M. D., Yee, H. K. C., & Ellingson, E. (2002), AJ 123, 1

Gobat, R. et al. (2018), Nature Astronomy 2, 239

González Delgado, R. M. & Cid Fernandes, R. (2010), MNRAS 403, 797

Grazian, A. et al. (2017), A&A 602, A18, A18

Grebel, E. K., Gallagher III, J. S., & Harbeck, D. (2003), AJ 125, 1926

Greig, B., Mesinger, A., Haiman, Z., & Simcoe, R. A. (2017), MNRAS 466, 4239

Griffith, E., Martini, P., & Conroy, C. (2019), MNRAS 484, 562

Grogin, N. A. et al. (2011), ApJS 197, 35, 35

Groves, B. A., Heckman, T. M., & Kauffmann, G. (2006), MNRAS 371, 1559

Gunn, J. E. & Peterson, B. A. (1965), ApJ 142, 1633

Guo, Q. et al. (2016), MNRAS 461, 3457

Hafen, Z. et al. (2017), MNRAS 469, 2292

Harrison, C. D. et al. (2011), MNRAS 413, 1036

Hayward, C. C. & Hopkins, P. F. (2017), MNRAS 465, 1682

Heckman, T. M., Sembach, K. R., et al. (2001), ApJ 554, 1021

Heckman, T. M., Borthakur, S., et al. (2011), ApJ 730, 5, 5

Henry, A., Scarlata, C., Martin, C. L., & Erb, D. (2015), ApJ 809, 19, 19

Hopkins, P. F., Kereš, D., et al. (2014), MNRAS 445, 581

Hopkins, P. F., Quataert, E., & Murray, N. (2012), MNRAS 421, 3522

Hopkins, P. F., Somerville, R. S., et al. (2006), ApJ 652, 864

Hung, C.-L. et al. (2019), MNRAS 482, 5125

Ishigaki, M. et al. (2018), ApJ 854, 73, 73

Izotov, Y. I., Orlitová, I., et al. (2016), Nature 529, 178

Izotov, Y. I., Schaerer, D., et al. (2018), MNRAS 474, 4514

Jian, H.-Y. et al. (2018), PASJ 70, S23, S23

Jones, T. A., Swinbank, A. M., Ellis, R. S., Richard, J., & Stark, D. P. (2010), MNRAS 404, 1247

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

Jones, T., Wang, X., et al. (2015), AJ 149, 107, 107

Jones, T. A., Ellis, R. S., Schenker, M. A., & Stark, D. P. (2013), ApJ 779, 52, 52

Jones, T., Stark, D. P., & Ellis, R. S. (2012), ApJ 751, 51, 51

– (2018), ApJ 863, 191, 191

Jørgensen, I. (1999), MNRAS 306, 607

Jørgensen, I., Chiboucas, K., Webb, K., & Woodrum, C. (2018), AJ 156, 224, 224

Jullo, E. et al. (2007), New Journal of Physics 9, 447

Kacprzak, G. G. et al. (2015), ApJ 802, L26, L26

Katz, N. & Gunn, J. E. (1991), ApJ 377, 365

Kawinwanichakij, L. et al. (2017), ApJ 847, 134, 134

Kennicutt Robert C., J. (1998), Annual Review of Astronomy and Astrophysics 36, 189

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. (2005), MNRAS 363, 2

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. (2001), ApJ 556, 121

Kewley, L. J. & Ellison, S. L. (2008), ApJ 681, 1183

Kewley, L. J., Geller, M. J., & Jansen, R. A. (2004), AJ 127, 2002

Kewley, L. J., Zahid, H. J., et al. (2015), ApJL 812, L20, L20

Khaire, V., Srianand, R., Choudhury, T. R., & Gaikwad, P. (2016), MNRAS 457, 4051

Kimm, T. & Cen, R. (2014), ApJ 788, 121, 121

Kirby, E. N., Rizzi, L., et al. (2017), ApJ 834, 9, 9

Kirby, E. N., Cohen, J. G., et al. (2013), ApJ 779, 102, 102

Kirby, E. N., Simon, J. D., & Cohen, J. G. (2015), ApJ 810, 56, 56

Kochanek, C. S. et al. (2012), ApJS 200, 8, 8

Koleva, M., Prugniel, P., Ocvirk, P., Le Borgne, D., & Soubiran, C. (2008), MNRAS 385, 1998

Koyama, Y. et al. (2013), MNRAS 434, 423

Kriek, M., van Dokkum, P. G., Franx, M., Illingworth, G. D., & Magee, D. K. (2009), ApJL 705, L71

Kroupa, P. (2001), MNRAS 322, 231

Krumholz, M. R., Dekel, A., & McKee, C. F. (2012), ApJ 745, 69, 69

Kuhlen, M. & Faucher-Giguère, C.-A. (2012), MNRAS 423, 862

Kurucz, R. (1993), SYNTHE Spectrum Synthesis Programs and Line Data. Kurucz CD-ROM No. 18. Cambridge 18

Lagos, C. d. P., Lacey, C. G., & Baugh, C. M. (2013), MNRAS 436, 1787

Lam, D. et al. (2019), arXiv e-prints, arXiv:1902.02786, arXiv:1902.02786

Larkin, J. et al. (2006), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Larson, R. B. (1974), MNRAS 169, 229

Law, D. R., Steidel, C. C., Erb, D. K., et al. (2009), ApJ 697, 2057

Law, D. R., Steidel, C. C., Shapley, A. E., et al. (2012), ApJ 759, 29, 29

Le Borgne, D. et al. (2004), A&A 425, 881

Le Borgne, J.-F. et al. (2003), A&A 402, 433

Lee, B. et al. (2013), ApJ 774, 47, 47

Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., Richard, J., et al. (2016), ApJ 820, 84, 84

Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. (2016), ApJ 831, 152, 152

Leethochawalit, N., Kirby, E. N., Moran, S. M., Ellis, R. S., & Treu, T. (2018), ApJ 856, 15, 15

Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. (2013), A&A 553, A106, A106

Leitet, E., Bergvall, N., Piskunov, N., & Andersson, B. .-.-G. (2011), A&A 532, A107, A107

Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. (2016), ApJ 823, 64, 64

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

Lian, J. et al. (2018), MNRAS 474, 1143

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. (2013), ApJ 772, 119, 119

Limousin, M. et al. (2012), A&A 544, A71, A71

Liu, F. S. et al. (2018), ApJ 860, 60, 60

Livermore, R. C., Jones, T. A., et al. (2015), ArXiv e-prints

Livermore, R. C., Jones, T., et al. (2012), MNRAS 427, 688

Lu, Y., Benson, A., et al. (2017), ApJ 846, 66, 66

Lu, Y., Mo, H. J., Weinberg, M. D., & Katz, N. (2011), MNRAS 416, 1949

Lu, Y., Blanc, G. A., & Benson, A. (2015), ApJ 808, 129, 129

Lu, Y., Wechsler, R. H., et al. (2014), ApJ 795, 123, 123

Ma, X. (2018), PhD thesis, TAPIR, MC 350-17, California Institute of Technology, Pasadena, CA 91125, USA

Ma, X., Hopkins, P. F., Feldmann, R., et al. (2017), MNRAS 466, 4780

Ma, X., Hopkins, P. F., Kasen, D., et al. (2016), MNRAS 459, 3614

Ma, X., Kasen, D., et al. (2015), MNRAS 453, 960

Maciel, W. J., Costa, R. D. D., & Uchida, M. M. M. (2003), A&A 397, 667

Madau, P. & Dickinson, M. (2014), Annual Review of Astronomy and Astrophysics 52, 415

Maiolino, R. & Mannucci, F. (2019), Astronomy and Astrophysics Review 27, 3, 3

Maiolino, R., Nagao, T., et al. (2008), A&A 488, 463

Malavasi, N. et al. (2017), MNRAS 470, 1274

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

Maoz, D., Mannucci, F., & Brandt, T. D. (2012), MNRAS 426, 3282

Maraston, C. (2005), MNRAS 362, 799

Marigo, P. & Girardi, L. (2007), A&A 469, 239

Marigo, P., Girardi, L., et al. (2008), A&A 482, 883

Markwardt, C. (2012), Astrophysics Source Code Library

Martin, C. L. (2005), ApJ 621, 227

Mason, C. A., Fontana, A., et al. (2019), MNRAS 485, 3947

Mason, C. A., Treu, T., et al. (2018), ApJ 856, 2, 2

Mathews, W. G. & Baker, J. C. (1971), ApJ 170, 241

McAlpine, S. et al. (2016), Astronomy and Computing 15, 72

McClure, R. D. & van den Bergh, S. (1968), AJ 73, 1008

McGreer, I. D., Mesinger, A., & D'Odorico, V. (2015), MNRAS 447, 499

McLeod, D. J., McLure, R. J., & Dunlop, J. S. (2016), MNRAS 459, 3812

Mendez, A. J. et al. (2011), ApJ 736, 110, 110

Mentz, J. J. et al. (2016), MNRAS 463, 2819

Meurer, G. R., Heckman, T. M., & Calzetti, D. (1999), ApJ 521, 64

Mieda, E. et al. (2014), PASP 126, 250

Miralda-Escude, J. & Ostriker, J. P. (1990), ApJ 350, 1

Miralda-Escudé, J. (1998), ApJ 501, 15

Mitra, S., Choudhury, T. R., & Ferrara, A. (2015), MNRAS 454, L76

Moran, S. M., Ellis, R. S., Treu, T., Salim, S., et al. (2006), ApJL 641, L97

Moran, S. M., Miller, N., Treu, T., Ellis, R. S., & Smith, G. P. (2007), ApJ 659, 1138

Moran, S. M., Ellis, R. S., Treu, T., Smail, I., et al. (2005), ApJ 634, 977

Moran, S. M., Ellis, R. S., Treu, T., Smith, G. P., et al. (2007), ApJ 671, 1503

Morishita, T. et al. (2017), ApJ 835, 254, 254

Mostardi, R. E. et al. (2015), ApJ 810, 107, 107

Moustakas, J. et al. (2011), ArXiv e-prints

Muñoz-Mateos, J. C. et al. (2007), ApJ 658, 1006

Muratov, A. L. et al. (2015), MNRAS 454, 2691

Murray, N., Quataert, E., & Thompson, T. A. (2005), ApJ 618, 569

Muzzin, A., Marchesini, D., et al. (2013), ApJ 777, 18, 18

Muzzin, A., Wilson, G., et al. (2012), ApJ 746, 188, 188

Naab, T. & Ostriker, J. P. (2017), Annual Review of Astronomy and Astrophysics 55, 59

Naidu, R. P., Forrest, B., Oesch, P. A., Tran, K.-V. H., & Holden, B. P. (2018), MNRAS 478, 791

Naiman, J. P. et al. (2018), MNRAS 477, 1206

Nakajima, K. & Ouchi, M. (2014), MNRAS 442, 900

Navarro, J. F. & White, S. D. M. (1994), MNRAS 267, 401

Nelson, D. et al. (2015), ArXiv e-prints

Nelson, E. J. et al. (2012), ApJ 747, L28, L28

Newman, S. F., Buschkamp, P., et al. (2014), ApJ 781, 21, 21

Newman, S. F., Genzel, R., Förster Schreiber, N. M., et al. (2013), ApJ 767, 104, 104

Newman, S. F., Genzel, R., Förster-Schreiber, N. M., et al. (2012), ApJ 761, 43, 43

Niemiec, A., Jullo, E., Giocoli, C., Limousin, M., & Jauzac, M. (2018), arXiv e-prints, arXiv:1811.04996, arXiv:1811.04996

Noeske, K. G. et al. (2007), ApJL 660, L47

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

Ocvirk, P., Pichon, C., Lançon, A., & Thiébaut, E. (2006), MNRAS 365, 46

Oesch, P. A. et al. (2015), ApJ 808, 104, 104

Okamoto, T., Nagashima, M., Lacey, C. G., & Frenk, C. S. (2017), MNRAS 464, 4866

Onodera, M., Carollo, C. M., Lilly, S., et al. (2016), ApJ 822, 42, 42

Onodera, M., Carollo, C. M., Renzini, A., et al. (2015), ApJ 808, 161, 161

Paardekooper, J.-P., Khochfar, S., & Dalla Vecchia, C. (2015), MNRAS 451, 2544

Pagel, B. E. J. (1997)

Panter, B., Jimenez, R., Heavens, A. F., & Charlot, S. (2008), MNRAS 391, 1117

Parravano, A., Hollenbach, D. J., & McKee, C. F. (2003), ApJ 584, 797

Patrício, V. et al. (2016), MNRAS 456, 4191

Pelló, R. et al. (1999), A&A 346, 359

Peng, Y., Maiolino, R., & Cochrane, R. (2015), Nature 521, 192

Peng, Y.-j. et al. (2010), ApJ 721, 193

Pettini, M. & Pagel, B. E. J. (2004), MNRAS 348, L59

Pilkington, K. et al. (2012), A&A 540, A56, A56

Pillepich, A. et al. (2018), MNRAS 473, 4077

Pilyugin, L. S. (2003), A&A 399, 1003

Planck Collaboration, Adam, R., et al. (2016), A&A 596, A108, A108

Planck Collaboration, Ade, P. A. R., et al. (2016), A&A 594, A13, A13

Pohlen, M. & Trujillo, I. (2006), A&A 454, 759

Porter, L. A. et al. (2014), MNRAS 445, 3092

Prugniel, P., Soubiran, C., Koleva, M., & Le Borgne, D. (2007), ArXiv Astrophysics e-prints

Queyrel, J. et al. (2012), A&A 539, A93, A93

Quider, A. M., Pettini, M., Shapley, A. E., & Steidel, C. C. (2009), MNRAS 398, 1263

Reddy, N. A., Steidel, C. C., Pettini, M., Bogosavljević, M., & Shapley, A. E. (2016), ApJ 828, 108, 108

Richard, J., Jones, T., et al. (2011), MNRAS 413, 643

Richard, J., Smith, G. P., et al. (2010), MNRAS 404, 325

Richards, G. T. et al. (2006), AJ 131, 2766

Rivera-Thorsen, T. E. et al. (2015), ApJ 805, 14, 14

Roberts-Borsani, G. W. et al. (2016), ApJ 823, 143, 143

Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. (2015), ApJ 802, L19, L19

Roškar, R. et al. (2008), ApJ 675, L65

Rubin, K. H. R. et al. (2014), ApJ 794, 156, 156

Ruiz-Lara, T. et al. (2018), A&A 617, A18, A18

Rzepecki, J., Lombardi, M., Rosati, P., Bignamini, A., & Tozzi, P. (2007), A&A 471, 743

Salpeter, E. E. (1955), ApJ 121, 161

Salvaterra, R., Haardt, F., & Ferrara, A. (2005), MNRAS 362, L50

Sánchez-Blázquez, P. et al. (2006), MNRAS 371, 703

Sanders, R. L. et al. (2015), ApJ 799, 138, 138

Saracco, P. et al. (2019), MNRAS 484, 2281

Schawinski, K. et al. (2014), MNRAS 440, 889

Schaye, J. et al. (2015), MNRAS 446, 521

Schenker, M. A., Robertson, B. E., et al. (2013), ApJ 768, 196, 196

Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. (2014), ApJ 795, 20, 20

Schiavon, R. P. et al. (2006), ApJL 651, L93

Schinnerer, E. et al. (2016), ApJ 833, 112, 112

Schmidt, M. (1965), ApJ 141, 1295

Schreiber, C. et al. (2015), A&A 575, A74, A74

Shapley, A. E., Reddy, N. A., et al. (2015), ApJ 801, 88, 88

Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. (2003), ApJ 588, 65

Shapley, A. E., Steidel, C. C., Strom, A. L., et al. (2016), ApJ 826, L24, L24

Sharples, R. et al. (2013), The Messenger 151, 21

Siana, B. et al. (2015), ApJ 804, 17, 17

Smit, R. et al. (2014), ApJ 784, 58, 58

Smith, G. P. et al. (2005), MNRAS 359, 417

Somerville, R. S., Gilmore, R. C., Primack, J. R., & Domınguez, A. (2012), MNRAS 423, 1992

Somerville, R. S. & Davé, R. (2015), ARA&A 53, 51

Sommariva, V. et al. (2012), A&A 539, A136, A136

Spitoni, E., Calura, F., Matteucci, F., & Recchi, S. (2010), A&A 514, A73, A73

Stark, D. P., Auger, M., et al. (2013), MNRAS 436, 1040

Stark, D. P., Swinbank, A. M., et al. (2008), Nature 455, 775

Stark, D. P., Walth, G., et al. (2015), MNRAS 454, 1393

Stark, D. P., Ellis, R. S., Bunker, A., et al. (2009), ApJ 697, 1493

Stark, D. P., Ellis, R. S., Charlot, S., et al. (2017), MNRAS 464, 469

Steidel, C. C., Rudie, G. C., et al. (2014), ApJ 795, 165, 165

Steidel, C. C., Erb, D. K., et al. (2010), ApJ 717, 289

Steidel, C. C., Strom, A. L., et al. (2016), ApJ 826, 159, 159

Stinson, G. S. et al. (2013), MNRAS 428, 129

Stott, J. P. et al. (2014), MNRAS 443, 2695

Strom, A. L. et al. (2017), ApJ 836, 164, 164

Sugahara, Y. et al. (2017), ApJ 850, 51, 51

Swinbank, A. M., Sobral, D., et al. (2012), MNRAS 426, 935

Swinbank, A. M., Webb, T. M., et al. (2009), MNRAS 400, 1121

Sybilska, A., Kuntschner, H., et al. (2018), MNRAS 476, 4501

Sybilska, A., Lisker, T., et al. (2017), MNRAS 470, 815

Tacconi, L. J. et al. (2018), ApJ 853, 179, 179

Taylor, P. & Kobayashi, C. (2016), MNRAS 463, 2465

Trager, S. C. & Somerville, R. S. (2009), MNRAS 395, 608

Trainor, R. F., Steidel, C. C., Strom, A. L., & Rudie, G. C. (2015), ApJ 809, 89, 89

Tremonti, C. A. et al. (2004), ApJ 613, 898

Treu, T. et al. (2003), ApJ 591, 53

Troncoso, P. et al. (2014), A&A 563, A58, A58

Trussler, J. et al. (2018), arXiv e-prints, arXiv:1811.09283, arXiv:1811.09283

van den Bosch, F. C. et al. (2008), MNRAS 387, 79

Vanzella, E., de Barros, S., Castellano, M., et al. (2015), A&A 576, A116, A116

Vanzella, E., de Barros, S., Vasei, K., et al. (2016), ApJ 825, 41, 41

Vasei, K. et al. (2016), ApJ 831, 38, 38

Vazdekis, A., Coelho, P., et al. (2015), MNRAS 449, 1177

Vazdekis, A., Sánchez-Blázquez, P., et al. (2010), MNRAS 404, 1639

Vogelsberger, M. et al. (2014), Nature 509, 177

Walcher, C. J., Coelho, P., Gallazzi, A., & Charlot, S. (2009), MNRAS 398, L44

Wang, X. et al. (2017), ApJ 837, 89, 89

Weinberg, D. H., Miralda-Escudé, J., Hernquist, L., & Katz, N. (1997), ApJ 490, 564

Werk, J. K. et al. (2014), ApJ 792, 8, 8

Weymann, R. (1967), ApJ 147, 887

Wilkinson, D. M., Maraston, C., Goddard, D., Thomas, D., & Parikh, T. (2017), MNRAS 472, 4297

Wise, J. H. et al. (2014), MNRAS 442, 2560

Wisnioski, E., Förster Schreiber, N. M., et al. (2015), ApJ 799, 209, 209

Wisnioski, E., Glazebrook, K., et al. (2011), MNRAS 417, 2601

Wizinowich, P. L. et al. (2006), PASP 118, 297

Worthey, G., Faber, S. M., Gonzalez, J. J., & Burstein, D. (1994), ApJS 94, 687

Wright, S. A., Larkin, J. E., Graham, J. R., & Ma, C.-P. (2010), ApJ 711, 1291

Wu, P.-F., Zahid, H. J., Hwang, H. S., & Geller, M. J. (2017), MNRAS 468, 1881

Wuyts, S. et al. (2011), ApJ 742, 96, 96

Yabe, K. et al. (2012), PASJ 64, 60, 60

Yang, C.-C. & Krumholz, M. (2012), ApJ 758, 48, 48

Yates, R. M., Kauffmann, G., & Guo, Q. (2012), MNRAS 422, 215

Yuan, T.-T., Kewley, L. J., & Rich, J. (2013), ApJ 767, 106, 106

Yuan, T.-T., Kewley, L. J., Swinbank, A. M., Richard, J., & Livermore, R. C. (2011), ApJL 732, L14, L14

Zackrisson, E., Binggeli, C., et al. (2017), ApJ 836, 78, 78

Zackrisson, E., Inoue, A. K., & Jensen, H. (2013), ApJ 777, 39, 39

Zahid, H. J., Dima, G. I., et al. (2014), ApJ 791, 130, 130

Zahid, H. J., Geller, M. J., Fabricant, D. G., & Hwang, H. S. (2016), ApJ 832, 203, 203

Zahid, H. J., Geller, M. J., Kewley, L. J., et al. (2013), ApJ 771, L19, L19

Zitrin, A., Broadhurst, T., Coe, D., et al. (2011), MNRAS 413, 1753

Zitrin, A., Broadhurst, T., Umetsu, K., et al. (2009), MNRAS 396, 1985

Zitrin, A. et al. (2015a), ApJ 801, 44, 44

Zitrin, A. et al. (2015b), ApJ 801, 44, 44